

UNIVERSITEIT UTRECHT

TNO

MASTER THESIS

**Expressive behaviour in robot-child
interaction**

Author:

Myrthe TIELMAN

3345343

Supervisors UU:

Prof. Dr. John-Jules MEYER

Prof. Dr. Michael MOORTGAT

Supervisors TNO:

Dr. Rosemarijn LOOIJJE

Prof. Dr. Mark NEERINCX

*A thesis submitted in fulfilment of the requirements
of 30 ECTS for the degree of Master of Science.*

July 2013

“At bottom, robotics is about us. It is the discipline of emulating our lives, of wondering how we work.”

Rod Grupen, 2008

UNIVERSITEIT UTRECHT

Abstract

Faculteit Geesteswetenschappen

Departement Wijsbegeerte

Master of Science

Expressive behaviour in robot-child interaction

by Myrthe TIELMAN

The subject of this thesis is expressive behaviour, those non-linguistics expressions of mental states which are used by people in interactions. Based on previous research and technical constraints a model for expressive behaviour for the Nao robot was built and implemented. This model adapts expressive behaviour of the robot to the emotions of the child it interacts with, as well as to emotionally relevant occurrences such as winning a game. An experiment with 18 children (mean age 8.89) and two Nao robots was done to evaluate this model and study the opinions about and (interaction) behaviour of children. In a within-subjects design the independent variable was if the robot displayed expressive behaviour during the interaction, which consisted of a short dialogue and a quiz game. The dependent variables, namely the opinions and behaviour of the children, were measured through questionnaires and video analysis. The feedback of the children in the questionnaires suggests that emotional expressions through the voice are less suitable as they decrease intelligibility, but that showing emotion through movement is considered a very positive trait for a robot. The results further indicate that children react more expressively and more positive to an expressive robot than to a non-expressive robot. From their positive reactions we can conclude that children enjoy interacting with an expressive robot more than with a non-expressive robot...

Acknowledgements

This thesis would not have become what it is today without the assistance I've had. I'd therefore first like to thank my supervisors, Mark Neerincx, Rosemarijn Looije, John-Jules Meyer and Michael Moortgat for their help and input, and TNO for giving me the opportunity to do this research. A second thank goes to Bert Bierman, who has adapted the WoOz for me and whose assistance with the Nao robots has been invaluable. Thanks also goes to Koen Hindriks, who has assisted me with GOAL. Further, I'd like to thank all members of the ALIZ-e project for their input and their thoughts on my project. Also, thank goes to all the interns at TNO who have helped me, Iris van Dam in particular for her help with the many little things which made my experiment possible. Finally, thanks to the Daltonschool Lange Voren for their enthusiasm and letting me perform my experiment at their school, with their children...

Contents

Abstract	ii
Acknowledgements	iii
1 Introduction	1
1.1 Expressive Behaviour	1
1.2 Project	2
1.2.1 ALIZ-e	2
1.2.2 Nao	3
1.2.3 Situated Cognitive Engineering	3
1.3 Research methods	4
1.4 Relevance of the subject	5
1.5 Structure	5
2 Human-Human Interaction	7
2.1 Introduction	7
2.2 Emotion	7
2.2.1 Arousal, Valence, Stance	8
2.2.2 Expression - Physical	9
2.2.3 Expression - Verbal	10
2.2.4 Use	11
2.3 Gesture	14
2.3.1 Conscious Gesture	14
2.3.2 Spontaneous Gesture	14
2.3.3 Use	16
2.4 Children	16
2.4.1 Emotion	17
2.4.2 Gesture	17
2.5 Conclusion	17
3 Human-Robot Interaction	18
3.1 Introduction	18
3.2 Social Robotics	18
3.3 Previous Research	20
3.3.1 Emotion	20
3.3.2 Gesture	22

3.3.3	Children	23
3.3.4	Results & Difficulties	24
3.4	Expressive behaviour in Nao	25
3.4.1	Design choices	25
3.4.2	Emotion	26
3.4.3	Gesture	28
3.4.4	Personality	29
3.5	Conclusion	30
4	Architecture	31
4.1	Introduction	31
4.2	Model	31
4.2.1	Input	32
4.2.2	Internal Parameters	33
4.2.3	Behaviour Selection	33
4.2.4	Output	34
4.3	Dependencies	35
4.3.1	Valence Robot	35
4.3.2	Arousal Robot	37
4.3.3	Prosody voice	39
4.3.4	Eye colour	39
4.3.5	Trunk position	40
4.3.6	Head position	40
4.3.7	Gesture movement	41
4.3.8	Gesture size	41
4.4	Timing	42
4.5	Use-case	42
4.6	Conclusion	44
5	Implementation	45
5.1	Introduction	45
5.2	GOAL	45
5.3	Communication	46
5.3.1	Wizard of Oz	46
5.3.2	Environment	47
5.4	Program	47
5.4.1	Knowledge Base	48
5.4.2	Belief Base	49
5.4.3	Actions	49
5.4.4	Program Main	50
5.4.5	Events	50
5.5	Conclusion	51
6	Experiment Preparations	52
6.1	Introduction	52
6.2	Arousal and Valence	52
6.2.1	Behaviour characteristics	53

6.2.2	Testing Objectivity	55
6.2.2.1	Method	56
6.2.2.2	Discussion	58
6.2.2.3	Conclusion	59
6.3	Dialog	60
6.3.1	Gesturing	60
6.3.2	Emotional occurrences	60
6.4	Quiz	61
6.4.1	Questions	61
7	Experiment	62
7.1	Introduction	62
7.2	Experimental Method	63
7.2.1	Experimental Design	63
7.2.2	Participants	64
7.2.3	Measures	64
7.2.4	Materials	66
7.2.5	Procedure	68
7.3	Results	69
7.4	Discussion and Conclusions	76
8	Conclusion	79
8.1	Conclusion	79
8.2	Implications and Further Research	81
	Glossary	84
	References	87
	Appendix 1: Use Cases and Requirements	93
	Appendix 2: Expressive Behaviour	98
	Appendix 3: Eye colours	100
	Appendix 4: Use Case Quiz	101
	Appendix 5: GOAL code	103
	Appendix 6: Instructions Arousal/Valence input	124
	Appendix 7: Dialogs	128
	Appendix 8: Gestures	133
	Appendix 9: Quiz Questions	135
	Appendix 10: Questionnaires	140
	Appendix 11: Correlations	145

Chapter 1

Introduction

Robots have been around for some time now, but they have long only been instruments for professionals. For the last 10 years though, robots have been making steps out of the professional environment and into people's homes. One of the goals in current robotics research is to develop robots which can be used by people without any specific training, including elderly people and children. This goal comes with its own challenges, since such a robot should be designed in such a way that it is capable of interacting with people in an intuitive way. One of the most intuitive forms of interaction for people is natural human-like interaction, which is the reason that much work in robotics has aimed to develop a robot which is capable of interacting in the same way as humans (Breazeal, 2003a; Scheutz, Schemerhorn, Kramer, & Anderson, 2007; Severinson-Eklundh, Green, & Hüttenrauch, 2003). Human interaction, though, is something very difficult to model because of the many, sometimes unknown, factors which play a role. One of the aspects which is important is the use of non-linguistic expressive behaviour. Although the fields of emotion generation and gesture generation have been studied in some detail, studies which combine these behaviours in a single model are surprisingly sparse. In this thesis I aim to contribute to the development of a robot capable of natural interaction by providing insight into how a robot can use expressive behaviour and what effect this has on its interaction with children. In this introduction I will present the context in which this study was executed, my research question and an outline of the rest of the thesis.

1.1 Expressive Behaviour

The subject of this thesis is the role of expressive behaviour in robot-child interaction. The first question we should ask is; what exactly is meant with expressive behaviour? One can argue that all behaviour is expressive, since everything we do expresses something, be it intentional or not. In this thesis, though, the term *expressive behaviour* will be used in a limited way, focussing

on the non-linguistic expression of a mental state as used in one-on-one interactions. Because the term *mental state* is still quite vague and has several interpretations, I will focus on those behaviours which reflect an emotional state or support verbal expression. The term *emotion* will here be used to refer to those subjective states of feeling which involve arousal (either high or low), have a specific valence (ranging from very positive to very negative) and have characteristic forms of expression. The expressive behaviour which typically supports verbal language is gesture, the object-independent moving of arms and hands while talking.

The goal of this thesis can be split into two parts. The first part is to develop expressive behaviour for the Nao robot. The research question corresponding to this goal is:

- How can we make the Nao use expressive behaviour in robot-child interaction?

The second goal of this thesis is to see the effect of expressive behaviour by the Nao robot in robot-child interaction. The research question corresponding to this goal is:

- What is the effect of the Nao using the model for expressive behaviour in interaction with children?

1.2 Project

The research presented in this thesis was done within the context of the EU project ALIZ-E. This section will introduce the ALIZ-E project and the specific demands this project placed on this thesis.

1.2.1 ALIZ-e

ALIZ-E is a project funded by the EU which started in 2010 and will last four years, ending in August 2014. This project has as its aim to develop methods to develop a robot which can interact with children for a longer, possibly discontinuous period of time. The scenario for which this robot is developed is a hospital stay for children with diabetes where the robot can be used to educate and entertain the child. Because of the complexity of the project, several different goals have been formulated. These goals focus on the different aspects of the research, such as long-term interaction, robot-child interaction, robustness of the interaction, real-world evaluation, the role of emotion and effect in robot-human communication, machine learning and adaptation and cloud computing. All of these subjects together should contribute to the development of a robot which is capable of a long-term any-depth affective interaction with children. Because of

the scale and the interdisciplinary aspect of this project, different European organizations are involved. These are the University of Plymouth (UK), the Deutsches Forschungszentrum für Künstliche Intelligenz (Germany), the Vrije Universiteit Brussel (Belgium), TNO (Netherlands), Imperial College London (UK), the University of Hertfordshire (UK), the Fondazione Centro San Raffaele del Monte Tabor (Italy), the National Research Council (Italy) and Gostai (France).¹

1.2.2 Nao

The robot used in the ALIZ-E project is Nao, developed by Aldebaran Robotics. (Figure 1.1) The Nao is a 57 cm tall humanoid robot. It has 25 degrees of freedom with joints at the head, shoulder, elbow, wrist, hip, knee and ankle. It has cameras, microphones and tactile sensors. Nao does not have any movable facial features, so it cannot display any facial expressions other than with head position and the colour of the eyes.²



FIGURE 1.1: Nao robot

1.2.3 Situated Cognitive Engineering

sCET stands for situated Cognitive Engineering Tool and is a tool designed to follow a certain method for research in cognitive robotics. This tool is based on the situated Cognitive Engineering method, developed by Looije, Neerinx, and Hindriks (2012). This tool was used when writing this thesis to structure the process and I will therefore first describe the method in more detail.

Situated Cognitive Engineering (sCE) was developed as a contribution to formalizing a baseline method for research in social robotics. Due to the large amount of usage contexts, theories, methods and technologies used in this field it has proven difficult to follow a single method in development and research projects. The sCE method takes into account the diversity of development and aims to be useful to all kinds of research in social robotics. The sCE approach consists of several design stages. The first stage is the derive stage. In this stage the operational demands, the human factors knowledge and the technical constraints are analysed. This stage should identify constraints on the design. The second stage is the specify stage, in which use-cases and claims are established. The claims are hypotheses about expected outcomes, often referring back to human-factors knowledge to keep the link to the theory. They are valid when the upsides outweigh the downsides. Linked to claims are requirements on the system, if a claim

¹Source: website ALIZ-e : <http://www.aliz-e.org>

²Source: website Aldebaran Robotics: <http://www.aldebaran-robotics.com>

is valid this can be seen as a justification of the requirements linked to it. Also formulated in the specify stage are use-cases. Use-cases are used to put the requirements into context and organize them. They take the form of scenarios, stories about people using the technology in context. These use-cases specify how far the theory extends to other situations. The third stage of sCE is the build stage in which the system is developed. No constraints are adopted for the kind of engineering methodology used, but sCE does promote a component-based approach. The fourth stage is the evaluate stage, in which the claims are tested with experiments. The fifth and final stage is the revision stage, in which the results from the evaluations are used to refine the system.

There are several reasons for adopting the sCE methodology. The first is that it is already used within the Aliz-e project by TNO and it is practical to stick to one specific methodology within the project. By using this method it will be easier to integrate the different systems developed within ALIZ-e in the end. This is, however, not the only advantage of using sCE. The sCE methodology provides a structured way to design models for social robotics. By starting with the theory this method makes sure that the product is well-grounded with respect to operational demands, human-factor knowledge and technical constraints. The use of use-cases ensures that the theory developed stays connected to the context of use and by developing use-cases it becomes easier to keep in mind what the system should actually do. The claims can then be used to keep sight of how the system could be tested to see if it works as it should. The build and evaluate stages give insight into the practical possibilities of the system and if it actually works as predicted. The revision stage provides a way to incorporate these insights into an improved system.

1.3 Research methods

To answer the research questions posed above I will use several research methods, each related to the sCE methodology. In order to answer the first question, we first need to know how people use expressive behaviour. This question will be answered by a study of previous literature from the field of psychology. I will then turn to literature from the field of robotics. Much research has already been done implementing forms of expressive behaviour in robots and their effect on interaction. These reviews can be seen as part of the derive stage as they give information about the human-factors knowledge and technical constraints. Based on the literature reviewed a model for expressive behaviour in the Nao robot can be designed. Important in the design of this model will be to formalize the requirements based on use-cases, the design corresponds with the specify stage. This model will be implemented to see if we can realize the model in the Nao robot. In order to answer the second research question, an experiment will be done to study the effect of the Nao's expressive behaviour. These steps correspond to the build and evaluate stage

respectively. Finally, the results from the experiment will be analysed so conclusions about a revision of the model can be drawn.

1.4 Relevance of the subject

Much work has been done in recent years in the field of human-robot interaction. Researchers have implemented emotion in robot and have developed models to make robots gesture and both in many variations. In this thesis, I aim to combine the knowledge gathered from these studies and their psychological background into a model which determines the expressive behaviour for a Nao robot and study its effects on interactions with children. The process of deciding on appropriate expressive behaviour is very relevant to the field of artificial intelligence. In the field of AI there is no consensus on the answer to the question if AI should model human intelligence or not. Despite this, there is little controversy on the idea that if we wish for an AI to interact in a simple and intuitive way with people, human interaction is what we wish to model. Many studies have already shown that specifics of a robots behaviour can have a great impact on their interaction with people. Chidambaram, Chiang, and Mutlu (2012) for instance, show that a robot displaying motivational non-verbal behaviour has a much greater impact on motivation than a robot which does not. This is an example of how robot behaviour can be used to improve human-robot interaction.

The subject of my own Masters programme is Cognitive Artificial Intelligence. The focus of this programme lies on the study of the different aspects of human intelligence (the fields of cognitive psychology and linguistics), the ways in which we can model intelligence in artificial systems (traditional AI and logic) and the fundamental questions we can ask in these fields (philosophy). Because of the wide scope of this program it is nearly impossible to find a specific research topic which combines all five research fields. Despite this, the subject of this thesis fits the program quite well as it combines knowledge from cognitive psychology and some from linguistics to develop an artificial model for expressive behaviour which will be subject to an experimental validation.

1.5 Structure

The rest of this paper is structured as follows. The next chapter will focus on expressive behaviour in natural human interaction. I will first look at the role of emotion; which emotions do we express in interaction and in which way? The second section will be about gestures as support for verbal utterances. In the final section, I will look at child-specific research to see

if children use expressive behaviour in a specific way. The third chapter will focus on human-robot interaction. I will first give an overview of previous research done in the field of social robotics. The next step is to look at the Nao robot, its abilities and limitations for displaying expressive behaviour. The final step is to see how the Nao robot should use expressive behaviour in interaction with a child. When should which emotion be displayed, etc. The fourth chapter will present a model for the Nao which represents the way in which the robot should use and display emotional behaviour in interactions. In the fifth chapter I will present the way the model for the Nao was implemented and the design choices made. Chapter six will describe the final preparations and design choices for the experiment aimed to test the model. Chapter seven describes the experimental design used in detail and presents its results. The final chapter will present the conclusions and suggestions for further research.

Chapter 2

Human-Human Interaction

2.1 Introduction

This chapter will look at the role of expressive behaviour in human-human interaction. As mentioned in the Introduction, I will adopt the term *expressive behaviour* to refer to those behaviours which consist of the non-linguistic expression of a mental state as used in one-on-one interactions. In this chapter I will look at the expression of emotion and gestures used to support speech. I will finish this chapter with a section about research focused on children to see if they express themselves differently than adults. This chapter is by no means meant to give a complete overview of the research done in the fields of emotion and gesturing, to do this would result in a very bulky book at the least. Since this is obviously beyond the scope of this thesis, I will focus instead on those studies and results which are most relevant when we wish to implement expressive behaviour in a robot. This also means that I will only consider those kinds of expressive behaviour which could in theory be used by the Nao robot. The most notable consequence of this selection is that any behaviour expressed through facial features will not be considered.

2.2 Emotion

In the history of research on the subject there has been much debate about the exact definition of *emotion*. Although no one would dispute that an emotion is a *strong feeling*, the question what exactly those strong feelings are has proven to be very difficult to answer and there are many different definitions still used simultaneously. In this paper, I will define emotional states as those subjective states of feeling which involve arousal (either high or low), have a specific valence (ranging from very positive to very negative) and have characteristic forms of expression. Note that in some cases *emotion* is only used to refer to those states of feeling

where it is clear what caused it, *mood* is used to describe those states where the cause is less clear and *affect* is used as a term encompassing both forms of feeling. In this paper I will not make this distinction and use the term *emotion* for all affective states. Much research has been done into the modelling of different emotional states and their relations. Ekman (1972) proposed that there are six distinct basic emotions which each have their own facial expression which is universally the same, namely happiness, sadness, anger, fear, disgust and surprise. Another view is presented by Plutchik (2001), who also defined eight bipolar emotions, pairs of joy vs. sorrow, anger vs. fear, acceptance vs. disgust and surprise vs. expectancy. The difference with the work of Ekman is that Plutchik stresses the relations between the emotions in a three-dimensional model of emotion. In this paper I will adopt a view more closely related with Plutchik's work, where emotions are classified by independent characteristics.

2.2.1 Arousal, Valence, Stance

The concept of classifying emotions on the basis of different factors is not a new one. Back in the 19th century, Wundt (1897) presented three psychological principles for classifying emotions. These were the quality of the feelings, the intensity of the feelings and the form of occurrence. The quality of the feeling could then again be divided into the three affective directions of pleasurable vs. unpleasurable, exciting vs. depressing and straining vs. relaxing emotions. These factors provided a method to classify emotions separate from specific terms such as *happy* or *sad*. In the 1950s the debate about which factors would be sufficient and necessary to describe emotion was rekindled by Nowlis and Nowlis (1956). Their work presented four dimensions for classification, namely *level of activation*, *level of control*, *social orientation* and *hedonic tone*. In the following 20 years, this debate resulted in the view that emotion could be described by a set of 6 to 12 monopolar factors. In the 1970s, however, a study by J. A. Russell and Mehrabian (1977) showed that the three independent, bipolar factors of pleasure-displeasure, degree of arousal and dominance-submissiveness are both necessary and sufficient to classify emotional states. In this thesis I will adopt a model of emotion related to the one by J. A. Russell and Mehrabian (1977), namely that of *arousal*, *valence* and *stance*, as presented first by Breazeal (2002). *Arousal* refers to how arousing the emotion is, an example of an emotion with low arousal is boredom, for high arousal this is excitement. *Valence* refers to how positive the emotion is, sadness being an emotion with a very low valence, happiness one with a high valence. *Stance*, finally, refers to how approachable the thing causing the emotion is. Fear is an emotion with a very low stance while interest has a high stance.

2.2.2 Expression - Physical

In this section I will focus on the expression of emotion. There are many different ways in which emotions are expressed, I will first look at the noticeable physical ways of expressing emotion. The most obvious way to express emotion is by facial expression, which I will not consider in detail since the Nao robot is not capable of simulating this kind of expression. People also, however, express emotion with body postures which is something the Nao is capable of doing.

Body language

Considering the long history of research into emotional facial expressions, research into the expression of emotion through body posture specifically is surprisingly sparse. Interest in the role of body language only started to grow in the past decade, though some studies focusing on the relation between facial and bodily expression of emotion have been done in the 20th century. Ekman (1965) looked at the role of head (including facial) and body cues in the recognition of emotions. In his study he used the factors sleep-tension and pleasantness-unpleasantness to classify the emotions, which are equivalent to the factors *arousal* and *valence* respectively. The results showed that while head cues carry more information about valence than about arousal, body cues carry almost no information about valence, but do give some about arousal. De Meijer (1989) looked at how we recognize emotion from body movement. From the results, three bipolar factors of emotion are derived for which specific movement combinations can be seen as indicators. The first factor is rejection-acceptance, where the mover shows either readiness to react with force or to interact without violence. The second factor is withdrawal-approach, where withdrawal suggests an indication to move away from the situation and approach suggests that the mover intends to move forward. The third factor is preparation-defeatedness, where preparation conveys a reaction to a sudden, unexpected situation and defeatedness conveys an inability to interact and loss of energy. The specific movements associated with these factors are shown in Figure 2.1. This table also shows that trunk movement is a good indicator of the valence of the emotion, it is stretched for positive and bowed for negative emotions. In the past decade, several studies have been done into the role of specific body postures in the expression of emotions. Coulson (2004) has developed methods to generate static body postures and review which of these postures are representative of which expressed emotion. Results show that the emotions of anger, sadness and happiness were well recognized from posture, fear and surprise were recognized less and disgust was never recognized from static postures. One suggestion from the author was that the low rates for fear and surprise might be due to a greater influence of movement in the expression of these emotions through body language. One of the current fields of interest is the development of a methodology for research into emotions in body language, to set standards for the description of body language and the ways these are generated (M. Gross, Crane, & Fredrickson, 2010; Dael, Mortillaro, & Scherer, 2012). No research has been done

	Rejection	Acceptance	Withdrawal	Approach	Preparation	Defeatedness
Trunk stretched		X		X	X	
Trunk bowed	X		X			X
Arms opening		X		X	X	
Arms closing			X			
Vertical upward						
Vertical downward			X			X
Sagittal forward				X		
Sagittal backward			X		X	
Strong force	X		X			
Light force		X			X	
Fast velocity	X					
Slow velocity						X
Direct movement		X	X		X	
Indirect movement	X					

FIGURE 2.1: Movements corresponding to emotion factors, adapted from de Meijer (1989)

yet to specifically investigate the way in which way arousal, valence and stance are reflected in emotional body postures or in movement. With the research which has been done though, it should possible to make some inferences about what movements and postures would be suitable for each factor.

2.2.3 Expression - Verbal

When people think about the expression of emotion, physical expressions are usually the first thing which comes to mind. Much of what we feel, however, is not expressed by how we look but by how we talk. This section will describe the different ways in which we express emotion with our voice.

We express emotion not only with what we say but also with how we say it. Research into vocal expression of emotion goes back to the first half of the 20th century. Studies have shown that people can recognize emotion from vocal cues alone with an accuracy of about 50%, well above chance. Some progress has also been made in establishing which acoustic cues represent which emotions. The main factor which has been studied is the fundamental frequency, its level, range and contour. Other factors are the amplitude, perceived as intensity, the distribution of the energy, the perception of articulation and temporal factors. Banse and Scherer (1996) studied the recognition rates and the influence of several of the previously mentioned factors in emotional speech for 14 different emotions. Their findings on recognition show that the emotions of anger, boredom and interest were recognised best of the emotions tested, shame and disgust worst. The other emotions were around the 50% rate. The mean fundamental

frequency was highest for what the authors call *intense* emotions, hot anger, panic fear and elation and lowest for contempt and boredom. The other emotions scored in the middle range. The mean energy was again highest for despair, hot anger, panic fear and elation and lowest for shame and sadness. Speech rate data showed that sadness had a particularly low speech rate, hot anger, panic fear, elation and happiness showed an increased rate. These results are consistent with previous findings.

2.2.4 Use

The previous section describes how we express emotions with body posture, movement and vocally. The next question is how we use this expressiveness in interaction. Do we use the expression of emotion in a certain way in certain situations in interactions? This section will try to give an answer to this question.

Own emotion

The first situation in which we express emotion seems obvious, when we feel it. We have a happy expression when we feel happy, our voice sounds sad when we feel sad. The story is a bit more complicated though, how we express our emotions differs per person and per situation. One of the processes which influences how we express emotion is emotion regulation. Research into emotion regulation has seen much growth in the past decade. It is distinguished from emotion generation in that emotion regulation refers to the processes which influence which emotions we experience, when we experience them and how we express them (J. J. Gross, Sheppes, & Urry, 2011). Emotion regulation can be intrinsic, when we regulate our own emotion, or extrinsic when someone else regulates our emotions such as a parent calming a child. Sometimes emotion regulation is explicit, when the emotion is consciously regulated, but it can also be unconsciously, implicitly regulated. When and how we regulate our emotions depends on situation and the person doing the regulating. The distinction between emotion generation and regulation has been very useful in the different fields of psychology and there is neurological evidence that instructions to regulate emotion leads to activation in parts of the brain associated with cognitive control and changes in regions associated with emotion generation (Poldrack, Wagner, Ochsner, & Gross, 2008).

Another factor which influences how we express emotion is gender. Theories about gender differences in emotion do not only look at whether women experience specific emotions different than men. Instead, much work looks at the dissimilarities related to differences in emotion regulation and recognition. One aspect which is related both to the amount of felt emotions and to the regulation is emotional expression. Despite the controversies in emotion research,

there is a general consensus in the field that women express more emotion than men. Though less well accepted, it seems as if this greater expressiveness for women is not exclusively related to the actual emotions felt. It is, of course, very difficult to measure what a person actually feels and studies either use self-reports, which can be influenced by social norms, or physiological measures, of which the exact relation to emotions felt is still largely unclear. Despite these methodical issues, some progress has been made. Studies show that women report to feel stronger than men when presented with the same emotional stimuli. Studies with physiological measures suggest that men have stronger reactions than women, particularly for fear and anger, while women react stronger to sadness and disgust (Kring & Gordon, 1998). There is also the question whether these gender differences are biological or cultural. A review by Brody (1985) suggests that many differences between men and women in the expression and self-report of emotion are not solely innate, since these differences are not the same for children in different stages of development. When children grow up, girls increasingly repress the expression and recognition of emotions which are socially unacceptable such as anger, boys increasingly repress the expression and attribution of most emotions.

Finally, how we express emotion is influenced by the social situation. We express emotions differently depending on the company we have. Buck, Loscow, Murphy, and Costanzo (1992) studied the role of the emotion and the relationship with the other person present on the expressiveness of the emotion. The results were consistent with previous studies (Fridlund, 1991; Yarczower & Daruns, 1982; Kleck et al., 1976), showing that when the person with us is familiar and sharing the same emotion, we express the emotion stronger than when we are alone. When we are with a stranger and experiencing a, generally more negative, emotion, we inhibit the expression of the emotion.

Mimicking

Aside from expressing emotion when we feel it, we also express emotion when the person we are interacting with shows it. We mimic the expressions of the person we're interacting with, smile when they are happy, frown when they're angry. This mimicking is not a purely expressive action, it is also important for the interaction. Bavelas, Black, Charles, and Mullett (1986) suggest that motor mimicry is in the first place communicative and that it conveys the message that *I am like you, I feel as you do*, a message fundamental in our relationship with others. In this way, mimicry strengthens our social relationships. In a study by Hess and Blairy (2001), the relationship between mimicking facial expressions and emotional contagion and emotional communication was tested. Results show that all facial expressions shown (happiness, sadness and anger) were mimicked by the participants. An emotional contagion effect was also found, but there was no indication that mimicry and emotional contagion were related. Moreover, no

effect of mimicry was found on communication. This would suggest that mimicry is not related to the own emotion experienced and that mimicry is not communicative. The question of what the exact purpose of mimicry is, is therefore still open. Chartrand and Bargh (1999) suggest that the process behind mimicry is the perception-behaviour link, the perception of a certain behaviour interferes with our motor process, causing us to execute the same behaviour. This would mean that mimicry is a purely unconscious motor process. Some studies, though, have shown that incongruent mimicry also exists, in some specific situations we do not mimic or even show opposite behaviours. Results have shown that we mimic less or incongruently when we do not agree with people, do not like them or when they are outside our social group. (McHugo, Lanzetta, & Bush, 1991; Bourgeois & Hess, 2008; Likowski, Mahlberger, Seibt, Pauli, & Weyers, 2008).

Contagion

Emotional contagion is a process with similarities to mimicry. Mimicry is the copying of expressions, a behaviour of which the link with the actual emotion felt is not entirely clear. Emotional contagion on the other hand, is the process where a person's emotions will become more similar to those expressed by our interaction partners. Mimicry is about behaviour, contagion about feelings. These two processes are, however, not fully separated and it is not quite clear in which way they interact. McHugo, Lanzetta, Sullivan, Masters, and Englis (1985) studied the effect of observing expressions by president Reagan on their subjects. They found evidence for mimicry, but also for emotional contagion. Both autonomic indicators and self-report measures showed that people were influenced by the expressions shown. When the emotion was happiness, people felt happier, when it was sadness they felt sadder, etc. Interestingly, the amount of contagion following the self-reports was influenced by the subject's opinion of president Reagan while that following the autonomous measures was not. This indicates that emotional contagion is an automatic process which affects us even when we don't really want it to.

Social role

The expression of emotion in interaction does not only serve an expressional purpose, it also has a social role. In the past decade, the interest into the role of emotions in social relationships has grown. Research has shown that emotion is a very important aspect in inter-personal relationships. Butler et al. (2003) have shown that when people suppress their emotions during first interactions, it disrupts communication, has a negative influence on emotional experience and that people report an inhibition of relationship formation. Emotion suppression thus has a negative influence on the formation of social relationships. In a related study, Peters and Kashima (2007) present results showing that emotional social talk strengthens the bond between narrator and audience and that their bond with the subject of the emotional talk is determined

by the emotionality of the anecdotes. Anderson, Keltner, and John (2003) show that when two people in a relationship are emotionally similar, the relationships were more cohesive and longer-lasting. The results from these studies indicate that emotion sharing is a very important aspect in forming and maintaining social relationships.

2.3 Gesture

When people are speaking, they usually do not sit still. One of the kinds of movements we make when talking is gesturing. These gestures are not the same kind as people use in sign language, gesture accompanying speech is spontaneous and strongly coupled with what we are saying. These gestures serve as a way to emphasize or support what were saying. In this section, I will describe the different kinds of gestures we use when talking and review the research into when and how we use which kind of gesture.

2.3.1 Conscious Gesture

The first kind of gesture to describe is the conscious gesture. These are the gestures we knowingly make. Within the class of conscious gestures, we can distinguish two different kinds, emblematic and propositional gestures. Emblematic gestures are those gestures which are culturally specified in that they are learned, their interpretation can differ between cultures. A typical example of an emblematic gesture is the *thumbs up*. Although emblematic gestures are one of the first to come to mind, they are actually used very little compared to other gestures and do not necessarily correlate with speech. The other kind of gesture made consciously is the propositional gesture. Propositional gestures have a direct relationship to what is being said, such as pointing to an object while saying *that*, or measuring with your hands when saying *she was this tall*. Although propositional gestures occur more than emblematic gesture, they still do not make up the majority of gesture in speech (Cassell, 2000).

2.3.2 Spontaneous Gesture

The majority of gestures produced while speaking is spontaneous, people don't intentionally make them and are often not even aware of making them after they have done so. In this section I will first describe the different kinds of spontaneous gestures. Afterwards, I will take a look at how and when these gestures are used in interaction.

Kinds of gesture

McNeill (1992) describes four kinds of spontaneous gestures, namely iconic, metaphorical, beat and deictic gestures. Iconic gestures are those gestures which reference to concrete events. They are closely related to the semantic context of the utterance, representing the same scene as the speech they accompany. An example is the gesture of appearing to grab something and pull it back, accompanying the sentence *he grabs a big oak tree and he bends it way back*¹. The second kind of spontaneous gesture is the metaphoric gesture. Metaphoric gestures are pictorial like iconic gestures, but instead of representing a concrete event they present abstract ideas. The movement is metaphorical for the concept described. The third kind are beat gestures, these are not related to the specific semantics of an utterance. Beat gestures have a general form, namely the moving of the hand or fingers up and down or back and forth in a short, sharp movement. The production of beat gestures is related to the rhythm of speech. The final kind of gesture is deictic, a pointing gesture. Deictic gestures are not only used to point at a specific object, but can also be related to the semantic content of the speech in a metaphorical way.

Context of Use

Gestures are forms of movement and in this specific kind of movement different phases can be distinguished. Iconic and metaphorical gestures have three phases, a preparation in which the hand is moved from its original position, a stroke phase which contains the gesture content and a retraction phase. Beat gestures only have two phases (e.g. in/out, up/down), as do deictic gestures (point, retract). The distinction between the different phases is important when one looks at the timing of a gesture. Gestures do not just occur somewhere random during a sentence, they are synchronized with speech in a specific way. The stroke of the gesture co-occurs with the semantically or pragmatically linked part of speech. In the example mentioned above, the gesture starts with the word *oak*, this is the preparation phase as the words uttered at this time are unrelated to the gesture. Together with the utterance of the words *bends it way back* the stroke of the gesture is executed, this part of the sentence clearly does have a semantic relation with the specific gesture. The gesture preparation thus anticipates speech, while the gesture stroke is synchronized with it. Beat gestures do not have a semantic relation with speech, but are used to emphasize parts of speech. Beats can serve different pragmatic functions, such as signalling that a comment is unrelated to the plotline and maintaining a conversation as dyadic (Cassell, 2000).

Not all four kinds of spontaneous gestures occur equally often. Iconic and beat gestures are the most common, each making up approximately one third of gestures. About 5,5% of gestures is metaphoric, 3,5% deictic and 24% of gestures cannot be classified in one of the four categories. There are different types of clauses in speech. Narratives are subject to sequential constraints (the telling of a plot-line), while extranarratives are not (e.g. descriptions). If the spoken clause

¹Example taken from: *Hand and Mind*, by McNeill, page 25

is a narrative or extranarrative influences how much a certain type of gesture is used. Iconic and deictic gestures are found most in narrative clauses (+- 85% vs. 90% resp.), metaphoric gestures are found most in extranarrative clauses (+- 70%) and beat gestures occur equally often in both situations. Almost all gestures occur during the actual speaking, with only a few occurring during pauses (McNeill, 1992).

Gestures are universal in that they appear in every language, but the way people gesture is different for different cultures. Iconic gestures show great overlap cross-culturally due to their close relation with semantic content. Metaphoric gestures are also universal in that they appear across the globe, but they do show large differences in the kinds of gestures used for different abstract concepts.

2.3.3 Use

Gestures accompanying speech serve several related purposes. The most important role of gestures is that they tell us what the speaker finds relevant, as gestures tend to accompany those parts of speech the speaker wants to emphasize. This can also be seen from the observation that the majority of gestures occurs during the rheme of an utterance, those parts of the text spoken which are new or interesting. Moreover, gestures tend occur during intonational rises in speech and the stroke is usually synchronized with the pitch accent (Cassell, 2000). Another example of how gestures tell us what is important is a phrase such as *walking away*. If this phrase is accompanied by a walking gesture or an away gesture indicates which part of the utterance is most relevant. Related to this, gestures reveal something about the speaker to the listener. Many gestures are made in reference to the person making them, such as pointing away or to yourself. This kind of pointing can, for instance, show if a person feels close to something.

2.4 Children

In the previous two sections, we have seen how and when we express emotion and gesture while speaking. Almost all of the research described though, was done with adult subjects. Since we wish to come to a model for robot-child interaction, it is important to review if there are differences between adults and children in how they express themselves. In this section I will look at child-specific research into emotion and gesturing to get a clearer view of how appropriate earlier conclusions are for this thesis.

2.4.1 Emotion

There is a solid base of research into the development of emotion and emotional recognition in children. Although emotion is recognized best when presented through both facial features and voice, research has also shown that children recognize emotion presented through only one of these channels (Stifter & Fox, 1987; Hortacsu & Ekinci, 1992; Kolb, Wilson, & Taylor, 1992). Research into the recognition of body motion shows a slightly different pattern. Boone and Cunningham (1998) show that children as young as 4 years old can recognize sadness in movement and 5 year olds are also capable of recognizing happiness and fear. By the age of 8 years old, children are comparable to adults in the recognition of emotion from movement, suggesting a quick improvement between the age of 4/5 and 8.

2.4.2 Gesture

Children start to gesture at a very young age, before they start to speak. Guidetti (2002) show that children between 16 and 36 months already use emblematic gestures such as pointing, nodding and shaking their heads. Colletta, Pellenq, and Guidetti (2010) studied the changes in gestures accompanying narratives for 6 year olds, 10 year olds and adults. Results show that the use of gestures accompanying language increases with age. The adults gestured most, followed by the 10 year olds, followed by the 6 year olds. The increase in gestures was mostly an increase in non-representational gestures. The authors suggest that this specific trend in the development of gestures is linked to the development of discourse cohesion with age.

2.5 Conclusion

This chapter has presented an overview of research into expressive behaviour. The two different forms of expressive behaviour are emotion expression and gesturing. We have seen that emotion can be presented through physical and vocal cues. Gestures fall into two categories, those of conscious gestures and spontaneous gestures which are closely coupled with language. Finally, a look at child-specific research shows that emotion recognition and non-representational gesturing strongly develops during childhood. This indicates that children may have these skills to a lesser extent than adults.

Chapter 3

Human-Robot Interaction

3.1 Introduction

In the previous chapter I have looked at how expressive behaviour is used in human interactions. The goal of this thesis, however, is not just to see how people use this behaviour but also to find out how a robot can. Much work has already been done in developing robots which can express themselves. This chapter will first introduce the field of social robotics, what it encompasses and why it exists. After this introduction, I will look at previous research into making robots display expressive behaviour. At the end of this chapter I will turn to the Nao robot. With the previously studied literature in mind, I will present the theory of how the Nao robot should and can use expressive behaviour in interaction with children. This will provide the theoretical base of the architecture for the Nao robot.

3.2 Social Robotics

Breazeal (2003b) defines the term *social robot* as those robots to which people apply a social model in order to understand and interact with them. To these robots people attribute mental states such as intents, beliefs, feeling and desires to understand their complex behaviour. Breazeal (2003b) also identifies four different kinds of social robots. The first class are socially evocative robots, robots which are designed to encourage people to attribute human characteristics to them, but whose design goes no further than this. The second class are social interface robots, which use human-like social cues and communication modalities. The third class are socially receptive robots, who do not just interact in a human-like way, but also benefit from interactions with people themselves. The final class is that of sociable robots. These robots have their own internal states and goals and interact in a human-like way not only to the benefit of the interaction partner, but also to their own. In addition to these classes, Fong,

Nourbakhsh, and Dautenhahn (2003) define three other classes of social robots from different uses in the literature. The first is the class of socially situated robots, which are surrounded by an environment which they perceive and react to. The second class are socially embedded robots, which are situated in a social environment where they interact with people and robots, are structurally coupled with their social environment and are at least partially aware of interactional structures as used by people (Dautenhahn, Ogden, & Quick, 2002). The final class is that of socially intelligent robots, which show aspects of a human-like style of social intelligence, which is based on models of human cognition and social competence (Dautenhahn, 1998).

One of the first questions to be asked in the field of social robotics is; *why would we want social robots?* This question can be specified into two parts, dependent on the type of social robot. The first question, applicable to all types of social robots is: *why would we want robots to interact with humans in a human-like way?* The second question is *why would we want a robot with human-like internal state such as emotion, goals and motivations?* There is, of course, no one answer to either of these questions. Rather, there is a list of possible reasons which depend strongly on the goal of the robot. One of the most important reasons to want a robot to be sociable has to do with interface design. More and more, robots are meant to be used by people in a non-professional environment. For this reason, the robot should be as easy as possible to work with. The preferred situation is where the person using it does not need any training to interact optimally with the robot. This solution is very difficult to achieve, though. Many researchers agree with the conclusions of Reeves and Nass (1996), that the best approach to this problem is to have the robot interact in a human-like way, since this is the manner of interaction humans are most proficient in (Breazeal, 2003a; Scheutz et al., 2007; Severinson-Eklundh et al., 2003). In this context, the use of a robot capable of social interaction is that people need no training to interact with the robot in an optimal way. Another reason for social robots is to make them more enjoyable to interact with. The idea that we like to interact with social robots more than with robots without social skills stems from the fact that people attribute human-like qualities to robots to better understand and predict their behaviour (Reeves & Nass, 1996). Being able to predict the behaviour of others is useful in interaction and there is a general consensus in the field of social robotics that people prefer robots whose behaviour they can predict with social models over those who are socially unpredictable (Breazeal, 2003b). The two reasons mentioned for making robots interact as social beings are both from the perspective of the human interaction partner. Two additional reasons for developing social robots are from the robotics perspective. First there is the advantage of social learning processes such as imitation and emulation. Through these processes a robot can learn new tasks in an easy way. The second reason from the robot perspective is that the implementation of an internal state can help the robot in long term interaction, self-maintenance, decision making, etc. (Breazeal, 2003a).

3.3 Previous Research

In this section, I will review studies from the field of robotics in which the goal was to make the robot use some form of expressive behaviour. The first chapter of this paper already gave an insight into the complexity of making robots expressive in a human like way. Because of this complexity, most studies so far have focused on a specific aspect of expressive behaviour. For this reason I will review studies with respect to emotion displays and gesturing separately.

3.3.1 Emotion

Emotion display and use is one of the most studied fields in social robotics, but there is much variation between studies. The most important aspects where studies differ are the choice for specific emotions or parameters such as valence and arousal, the choice for a biologically inspired or functionally designed robot and the chosen medium of expression. In this section I will first discuss some studies to illustrate the different design choices made in previous research. I will then look at how the different media of expression have been used to express emotion.

Design

Almost all models of emotion for robots are to some extent biologically inspired in that they use an internal model of emotion. Some are, however, designed to fit emotion theory as much as possible while others take more liberties. Arkin, Fujita, Takagi, and Hasekawa (2003) have developed a model for dog-like behaviour based on ethological principles, where an emotionally based architecture is used. The model is used to learn new objects by associating their effects on internal motivational and emotional variables. These variables are then used to generate appropriate behaviour. Canamero and Fredslund (2001) use a stimulation model which is also strongly inspired by emotion theory. The level of stimulation influences the emotional state of the robot, e.g. a sustained high level of stimulation activates negative emotions while a sudden decrease following a high stimulation level. Schulte, Rosenberg, and Thrun (1999) take a more functional approach with a museum-tour robot. This robot uses a state machine for the emotional expression shown when asking people to step aside. At first it will be happy, when blocked neutral, when still blocked sad and when still blocked angry. A related study was done by Nourbakhsh (1999), who also developed a robot to be used for museum-tours, but instead of a simple state machine they employ a fuzzy state machine model. In this model, external events trigger the transitions between moods and multiple emotions can be blended to achieve smoother transitions between emotions. Hirth, Schmitz, and Berns (2011) have developed the UKL Emotion-based Control Architecture, an architecture for emotions in robots which can be used on different platforms and for different functionalities. This architecture takes into account

five functions of emotion, namely regulative, selective, expressive, motivational and rating functions. Hirth, Schmitz, and Berns (2012) successfully implemented this architecture for a robot playing tangram, influencing its motivation based on its emotions.

Body Language

One of the reasons research in robotics started looking into emotional body language is that psychological research has shown the relevance of body language on emotion recognition (Ekman, 1965; De Meijer, 1989; Coulson, 2004), another is that not all robotic platforms are capable of moving their facial features to display emotion. Beck, Canamero, and Bard (2010) generated several emotional key poses in the Nao robot and tested recognisability and the effect of head movement within poses. Their results show that people can recognize the emotional body language from the robot. The effect of head position was great, overall perceived arousal, valence and stance were higher when the head was tilted upwards, lower when the head was lowered. In a later study (Beck, Hiole, Mazel, & Canamero, 2010), the affective space for body movement for the Nao is further developed. The approach here is similar to that of Kismet (Breazeal, 2001) in that it depends on the gradual factors of arousal and valence. The poses developed with this affective space were tested on recognisability and results show that the poses were recognized above chance. Moreover, 30%/70% mixes of emotions were still recognized at the emotion being displayed at 70%, while 50%/50% mixes were hardest to recognize. This is in line with the expectations and shows that the gradual change from one emotional body pose to another is well recognized. Beck et al. (2011) validated the previous findings with children and Häring, Bee, and André (2011) give further evidence that emotional body poses are well-recognized by people in interactions.

Verbal

Studies aimed at producing emotional speech are quite recent. One of the main reasons for this is that the first aim in robot speech research has been to make them speak at all. Solid text-to-speech systems have simply not been around a very long time, and to generate emotional speech you first need neutral speech. In recent years though, the field has seen enough growth to give rise to a new body of research into the automatic production of emotional speech. Burkhardt and Stegmann (2009) review the work done up to that point. They point out that most speech-synthesis systems work not with the arousal, valence, stance model but with specific emotions. The reason for this is that valence and stance are difficult to relate to specific voice changes. Two main approaches can be distinguished based on the method of synthesis. One is to produce speech as natural as possible, but with low flexibility, the other is to have a high flexibility with less natural sounding speech. Tesser, Zovato, Nicolao, and Cosi (2010) describe a method to produce sad and happy speech. They use statistical methods to transform synthesis speech

based on a corpus of sad and happy recorded speech. The resulting speech was evaluated by an objective and a subjective method and both evaluations showed that the transformation was successful. The speech showed prosodic characteristics closer to the corpus and was recognized by people as sad or happy. Kessens, Neerincx, Looije, Kroes, and Bloothoof (2009) used emotional speech synthesis in combination with facial expressions for different roles of the iCat in interaction with children. Their results show that the emotional speech can be recognized correctly, but that care should be taken that the speech does not become too unintelligible. These results indicate that it is possible to adapt synthesised speech to be emotional in such a way that people notice the emotional content of the voice.

3.3.2 Gesture

Gesture is an important aspect of the liveliness of virtual agents and robots. Whether or not such a character moves while speaking makes a big difference in how much people like and trust it. For this reason, quite some research exist in the field of automatic gesture generation. One of the most important and challenging aspects of gesture generation is the synchronization with speech. Cassell, Vilhjlmsson, and Bickmore (2001) developed an animation toolkit which aimed to automatically generate gesture, eye gaze and other nonverbal behaviour from text in a correctly synchronized way. This system relies on a textual analysis to identify clause, theme, rheme, word newness, contrast, objects and actions. Once these are identified, appropriate non-verbal behaviour is suggested from a knowledge base, including beats, iconic gestures, contrast gestures, eyebrow behaviour, gaze direction and intonation. A special module then analyses all found appropriate gestures to come to a compatible gesture suggestion. An approach more based in linguistics was taken by Giorogolo (2010), who designed a semantics for the relation between iconic gestures representing spatial structures and speech. This semantics was implemented in an artificial agent and an experiment confirmed that the semantics is a good model of the process of gesture interpretation. In Aly and Tapus (2011) and Aly and Tapus (2012), the authors use Coupled Hidden Markov Models (CHMM) to find a coupling joint between audio and gesture (arm and head). This is then used to generate arm and head movement accompanying speech based on a database of coupled speech and gesture. The generated head and arm gestures achieve similarity scores of 62% and 55% respectively. Moreover, a subjective analysis shows that the gestures generated are quite similar to the ones from the databases.

Aside from generating gestures, some studies also look at the effect of this generated gesturing on interactions with people. Salem, Kopp, Wachsmuth, Rohlfing, and Joublin (2012) implemented a gesture generation system which was first developed for an artificial agent in a humanoid robot. The robot is able to generate the gestures quite well, but there is a problem with synchronization. In longer sentences, the gesturing tends to lag behind the speech output. In an

experiment, the authors study the effect of gesturing in interaction. Results show that a robot which gestures is perceived more positive than a robot which only speaks, even when the gestures are semantically incongruent with speech. Kim, Kum, Roh, You, and Lee (2012) studied the role of robot gesture on information transfer and perception of the robot. The results show that there was no effect of gesture on the information transfer. Surprisingly, results show that the robot with incongruent gesture behaviour was perceived significantly more positive than the robot with no gesture behaviour, which was in turn perceived more positive than the robot with congruent gesture behaviour. The authors suggest that these results might be explained with resource theory. If this is the case the robot displaying congruent gesture with speech would present a higher workload for the person listening because two channels have to be interpreted.

Based on the studies discussed here, we can conclude that much progress has been made in generating robot gestures. Timing and semantic congruity are important, but the exact influence of gestures on the interaction is still not quite clear.

3.3.3 Children

As seen in the previous chapter, children do not always interact in the same way as adults. This is the case for human-human interaction, but certainly also for human-robot interaction. Children are, for instance, more likely to attribute characteristics such as emotions and memory to robots when they are younger (Beran, Ramirez-Serrano, Kuzyk, Fior, & Nugent, 2011). This suggests that perceived animatism in robots decreases with age. Benítez Sandoval and Penalzoza (2012) held a survey among children from 6 to 13 years old about their expectations and wishes for robots. Their results showed almost 80% of children would like to see robots as educators. When asked if they would like to see robots in museums, about 10% wanted to see robots as part of the exhibition while about 80% would like to see the robot interacting with people. This shows that children are very open for interaction with robots and that they would like robots to interact with them. Robots as museum guides or educators are two of the most explored kinds of social robots. One study which did not only look at robot educators in a clinical, short period experiment was done by Kanda, Sato, Saiwaki, and Ishiguro (2007). They conducted a long-term experiment in a class-room environment. Three design principles were incorporated to sustain long-term interaction, namely personalized behaviour, pseudo development and confiding in personal matters. The results show that those children which established a peer-type relationship with the robot kept interacting with it during the whole experiment. Children who did not consider the robot a friend stopped being interested in the robot after five to seven weeks. This shows that for long term interaction with children, it is important that the robot establishes a bond with their interaction partner to make sure they keep interest. Another study on robot educators was done by Saerbeck, Schut, Bartneck, and

Janse (2010). The focus of this study was the role of social empathic behaviour on a learning task. A robot was implemented with the role of a tutor which could give non-verbal feedback, guide the attention of the student and which could show empathic behaviour. The effects of this behaviour was studied comparing the social robot with one which did not have this social behaviour. The results show that social supportive behaviour from the robot has a significant positive influence on the learning performance. Also, participants which interacted with the social robot were more motivated to perform the task. This shows that social behaviour in robots can have an important influence on learning and motivation when interacting with children. A related study was done by Leite, Castellano, Pereira, Martinho, and Paiva (2012), who studied the role of empathy in robot-child interaction in a classroom environment. The robot was capable of recognizing the users affective state and reacting empathically. Results show that empathy facilitated interaction and makes the children like the robot more. They also suggest, however, that empathy behaviour should be carefully selected, since inappropriate behaviours can have an opposite effect. Shahid, Kraemer, and Swerts (2011) propose a method to evaluate child-robot interaction using comparison with child-child interaction. Children indicated that they had more fun when playing with the robot than with playing alone, but less than when playing with a friend. A perception test revealed that children are most expressive when playing with other children, a little less expressive when playing with a robot and less expressive still when playing alone. This study shows that although children like interacting with the robot, robot-child interaction is not a replacement for child-child interaction. Children like their friends more than the robot. Taking into account the results from these studies, we can conclude that robot-child interaction is a very promising field. Children like robots and a correctly designed robot can establish a long-term social bond with children. Social interaction between robots and children facilitates communication, learning and motivation.

3.3.4 Results & Difficulties

The previous section has aimed to give an overview of the current state of research into expressive behaviour by robots. Results from research into the expression of emotion through body posture and voice have generated emotional expressions and they have been found to be well-recognizable by people. One aspect of the expression of emotion which is less well studied is the effect of these expressions on interactions. Most studies into emotion for robots are motivated by research from psychology which have shown emotion to be beneficial in human interaction. Human-robot interaction is, however, not the same as human-human interaction and further validation of the exact effects of emotional expression by robots would be preferred.

Much research has been done into gesture generation based on text analysis. Different systems have been developed and validated with some success. The exact effects of congruent or incongruent gesture-speech production are, however, not yet clear and may very well depend on the task at hand. More research into the effect of robot gesturing on interaction is therefore needed to predict its effects. Finally, research with robots and children has been very promising. Children are very suited to work with social robots because they are ready to ascribe human characteristics to robots and because they like them. These factors make children motivated to work with robots, which contributes greatly to the use for robots for educational or motivational purposes.

3.4 Expressive behaviour in Nao

In the second chapter, I have given an insight into how humans display and use expressive behaviour in interactions. This chapter, I have turned to the field of robotics. In the previous sections, we have seen which reasons we have for making robots social, what work has already been done into making robot display expressive behaviour and what the practical demands of the chosen platform mean for the possibilities of implementing expressive behaviour. In this section, I will use the conclusions from previous works to formulate what kind of expressive behaviour the Nao robot should and could use in its interactions and in which way it should do so. I will first look at some design issues which are relevant when implementing expressive behaviour. Next, I will formulate the ways in which emotion and gesturing should be used. Finally, I will say something about the personality and role of the robot and in which way this influences its expressive behaviour.

3.4.1 Design choices

When implementing expressive behaviour into a robot, there are several design issues which play a role. One of the most important ones and the one I wish to discuss here has to do with the difference between, for instance, socially evocative and sociable robots. Socially evocative robots only display expressive behaviour to appear more human-like. Sociable robots, on the other hand, have internal representations of mental states and benefit from their expressive behaviour themselves. Fong et al. (2003) define this difference more precisely, specifying a class of biologically inspired robots and a class of functionally designed robots. Biologically inspired robots are those where the designers aim to create a robot which simulates social behaviour internally. Functionally designed robots, on the other hand, are robots where the design objective is to create a robot which outwardly appears to be socially intelligent. For the architecture for expressive behaviour presented in this paper, the choice is to start with the aim of a biologically inspired robot. This means that the robot will have internal representations of,

for instance, its emotions. A very simple scenario which can show the difference between these design options is that of interaction with a child. The goal is here to mirror the emotion of the child. A functionally designed robot will receive the input that the child is happy and therefore display happy behaviour. A biologically inspired robot, on the other hand, will receive the input that the child is happy, adapt its own emotional state to happy and display the behaviour fitting with its emotional state. The reason for this design choice is that when you wish to reason about appropriate expressive behaviour, several factors need to be taken into account. When reasoning with multiple, possibly interfering factors, it is helpful to have internal parameters and not just input to work with.

3.4.2 Emotion

The first question to be answered in this section is which kinds of expressions the Nao robot should use for expressing emotion in interaction with children. Secondly, I will consider which emotions the robot should show and finally I will formulate which factors should influence the emotions of the robot and in which way. As we have seen in the section about the Nao robot, it cannot display facial emotions. This means that emotional expressions have to be expressed through body pose, body movement, eye colour and voice. For body poses, I suggest using the poses developed by Beck, Hiole, et al. (2010) for the Nao robot, supplemented by poses developed by Aldebaran ¹. These poses have been developed for the same platform and have been validated. These factors make them very suitable for this model. The main problem with using body poses for emotion in an expressive robot is that we also want the robot to move its arms and head while talking and/or listening. Since a happy pose, for instance, has the robot raise its arms and head, it is very difficult to model gesture at the same time. For this reason there is a need for other ways to express emotion which might not interfere with gesturing. These can be body movement, head position, voice characteristics and eye colour. With respect to body movement, work by De Meijer (1989) shows that there is a complex relation between emotion and movement. The downside of his work is that the data was analysed using different parameters than arousal, valence and stance. Because we do wish to couple the expression of emotion to these parameters, I suggest using only trunk movement as this has a clear relation with the valence of the emotion involved and this does not interfere with gesturing. Work by Beck, Canamero, and Bard (2010) shows that head position has a strong influence on perceived arousal, valence and stance, the head goes up when these are high, down when these are low. These head positions can be used to express emotion in situations where the robot does need to gesture. Most current methods of generating emotional speech use statistical methods to generate their output and they mostly work with discrete emotions. These two factors make it

¹<http://www.aldebaran-robotics.com>

difficult to use their results as we wish for a model based on values of arousal and valence. Looking at the data, though, we can conclude that there seems to be a relation between arousal and fundamental frequency, mean intensity and speech rate (Banse & Scherer, 1996). When arousal is higher, so are these aspects of speech. The suggestion here is to use this observation and use voice only to express arousal, since the relation with valence is much more complex. Eye colour is an aspect which is not often used in human-robot interaction, but it is useful because eye colour can be changed during interaction without interfering with the other modalities. In this way eye colour can be used to express emotion constantly. Eye colour has been used previously for the Nao robot by Cohen, Looije, and Neerincx (2011), who used colours linked to emotions following the work of Kaya and Epps (2004). The results from this study can be used to come to a selection of colours linked to specific arousal values.

Now it is determined which ways of expressing emotion will be used in relation to the arousal and valence of the robot, the next step is to look at how these emotions will be expressed. First it is necessary to put some constraints on the values of arousal and valence. The first constraint has to do with the observation that children recognize emotional expressions less well than adults. For this reason, I suggest that a model aimed at children should have the values of emotion and valence change more than a model aimed at adults. This would mean that both the upper and lower bounds of the values are further apart and that rising or dropping values should move more. These adaptations would result in more exaggerated emotions, which is exactly what we want as these are easier to recognize. We should, of course, always be careful not to make the expressions too extreme, as we do wish them to be life-like. A second adaptation to be made is to make the robot have a generally positively valenced emotion. We wish the child to feel positively about the robot, so the robot should be mostly cheerful. For the same reason, we don't want the robot to be angry too quickly, so the combination of a high arousal and low valence should only be used in extreme circumstances. We don't want the robot to get angry when the child fails at something, only to get a little annoyed or disappointed if the child for instance keeps insulting the robot for a longer time. The values of both arousal and valence should have thresholds. Arousal has a lower and an upper threshold, lower to not become too boring, upper to not excite the child too much. Valence has a lower threshold to make sure that the robot does not become too depressing.

The final step is to look at how the robot should adapt its emotional expressive behaviour to the child and the situation. Input for the robot are the current arousal and valence of the child and the current task state. The first thing the robot should do with these values is mimic the child, or be influenced by them. People copy and are influenced by each other's emotional expressions in interactions and we wish for the robot to do the same. There should be some regulation, though, we do not wish a robot which was sad to suddenly become very excited because the

child is. Instead, we wish to adapt the arousal and the valence in the direction of the arousal and valence of the child. If the robot is sad, it will become happier when the child is happy. How much happier depends on the values for arousal and valence, the more they differ from the initial state of the robot, the more this will change. There are some thresholds involved in this mimicking process. If the child is extremely excited, the robot should have a calming influence and if the child is very sad the robot should have a cheering influence. For this reason there is a certain high level of arousal and low valence of the child where the robot will not simply mimic the values, but take into account that its own arousal cannot be too high or too low, nor can its own valence be too low. We do not wish the robot to only mimic the emotion of the child, there are also other factors which might influence the emotional state of the robot. Most of these will arise from the current task the robot is performing. For instance, when the robot has achieved one of its goals this will produce a rise in valence and arousal, when the robot fails to perform one of its task several times, arousal will rise and valence drop. One last important aspect of the robot's emotional behaviour is to select the right kind of movement and voice change, as there are some constraints. Since we wish for the robot to be able to gesture while talking, it cannot constantly display emotional body poses. For this reason, these poses will only be displayed for special occurrences during the task which are linked with specific strong emotions. An example is winning a game, the robot will display a happy pose, but then go back to the pose it started from with a slightly higher valence and arousal. Head and trunk movement, on the other hand, will remain for a longer time and will also be affected by small changes in emotional state. Thresholds for the changes in fundamental frequency, volume and speech rate are necessary to ensure that the robot remains understandable and pleasant to listen to.

3.4.3 Gesture

When developing a robot which should display natural expressive behaviour, gestures are very important. Although the first thing which comes to mind when hearing the word *gesture* are conscious gestures, we use spontaneous gestures much more often. For this reason I propose to only use one conscious gesture, namely pointing. More important for natural conversation, though, are spontaneous gestures. Beat gestures and iconic gestures should be used most, since these occur most often in human interaction. These gestures should be produced during speech and be related to the linguistic content of the utterances. Both semantics and pragmatics should be taken into account. Moreover, the gestures should be timed in such a way that they relate to speech correctly. There are roughly two ways to generate gestures when talking. The first is the easiest, which is to simply hard-code fitting gestures for every dialogue option. This method has no flexibility, but for tasks where every piece of possible dialogue is written in advance you don't need flexibility. It would be preferred, though, to have several slightly different gesture types for pieces of dialogue which occur more than once to avoid repetition.

The other approach is much more difficult but gives every flexibility, a version of this approach has been implemented by Cassell et al. (2001). This is to analyse a piece of text linguistically to generate the possible gestures for the semantic content including their timing. The next step is to evaluate if the gestures are incompatible with one another and select a combination of gestures which is feasible. Because the focus of this thesis lies on integrating the different kinds of non-verbal behaviour and not on gesture generation, this thesis will adopt the first approach, namely a hard-coded set of gestures, of which one will be chosen following research on how much specific kinds of gestures occur.

3.4.4 Personality

In this section, I will look at the influence of the personality and the role of the robot in its behaviour. One of the most important aspects of personality in this context is the notion of extravert and introvert personalities. In psychology, this notion is one of the five main traits which define a person's personality (Digman, 1990). Jung, Lim, Kwak, and Biocca (2012) studied the role of extravert and introvert personalities on human-robot interaction. They developed one introvert and one extrovert robot and studied how much these were liked by people with introvert and extrovert personalities. Their results show that introvert people felt more friendliness towards the introvert robot and liked it more. Extrovert people, on the other hand, felt more friendliness towards the extrovert robot and liked that one more. These results show that robot personality is important in human-robot interaction and that the effects of certain character traits in a robot differ per interaction partner. It would therefore be a good idea to adapt the personality of a robot to the personality of the user. The question which should now be asked is what the attributes are of personality traits, with which behaviour types can we implement personality? Borkenau and Liebler (1992) studied the effects of 45 different physical attributes on self-reported and perceived personality traits in people. Their results show that the greatest consensus was on the trait of extroversion, which indicates that this is the trait which is most easily observed correctly. This is important because if we wish to model personality in a robot, this personality should be understood well by the interaction partner. If a certain personality trait is not easily recognized by people it will be difficult to model this trait in a robot in such a way that it has effect. For this reason, I choose to only consider the notion of extroversion for the model of the Nao. Based on the work by Jung et al. (2012), we know that for this personality trait people tend to like a robot which has a personality similar to themselves. The final question, though, is which expressive behaviour characteristics influence personality perception. Borkenau and Liebler (1992) show that extrovert personalities are recognized by having a powerful voice, are easy to understand, have a friendly and self-assured expression, smile a lot, sit in a relaxed way, move quickly, move their hands and head a lot and touch themselves frequently. Introvert personalities, on the other hand, are soft-voiced,

speak haltingly, have an unconcerned expression, avoid looking at who they are talking to, don't swing their arms when walking and walk stiffly. Of course, not all these characteristics are applicable to the Nao robot, but we can make several generalisations. For an extrovert robot, the amount and the speed of gesturing is higher than for an introvert robot. They speak louder, look at their interaction partner more and are quicker to show positive emotions. All these factors can be incorporated in the model for the Nao by setting threshold values for these characteristics which are higher when interacting with introvert children than when interacting with extrovert children.

3.5 Conclusion

This chapter has focused on human-robot interaction. I have first given an overview of relevant research in the fields of emotion and gesture models for robots. This review shows which progress has been made, but also where the difficulties still lie. In the following sections I have turned to combining the conclusions from human-robot research and human-human research. I have presented the design choices for a model for the Nao robot with respect to emotion, gesture and the role of personality. In the next chapter, an architecture will be described which further formalizes the ideas presented in the previous section.

Chapter 4

Architecture

4.1 Introduction

In the previous chapters I have given an overview of the role of expressive behaviour in human interaction and presented an answer to the question how expressive behaviour should be used by the Nao robot in interactions with children. In this chapter I will formalize the constraints presented in section 3.4 in order to arrive at a model for an architecture usable when implementing expressive behaviour in the Nao robot. The goal of this chapter is to arrive at a model which does not need further adaptations in order to be implementable.

4.2 Model

The model to be presented in this chapter was developed in several stages. The first stage was to formulate which expressive behaviour the robot should show in which situation based on the human factors knowledge and technical constraints. This corresponds to the derive stage in the sCE method and was presented in the previous two chapters. The next step is the specify stage in which several context independent use-cases are formulated for the expressive behaviour of the robot. Different use-cases were formalised for the ways in which the robot should adapt its behaviour to the environment and are based on section 4 of the previous chapter. The first use-case describes the situation in which the robot mimics the emotion of the child. This is something the robot should do constantly, the exceptions being use case two and three, which deal with the child's valence being too low or his arousal being too high. The other external factor which influences the robot's emotion is task state and an example is discussed in use-case four. Use-cases five and six specify a situation in which the robot shows its emotional state while talking. Finally, use-case seven puts into context how the robot takes into account its extroversion, which is based on that of the child. The pre-conditions, post-conditions and

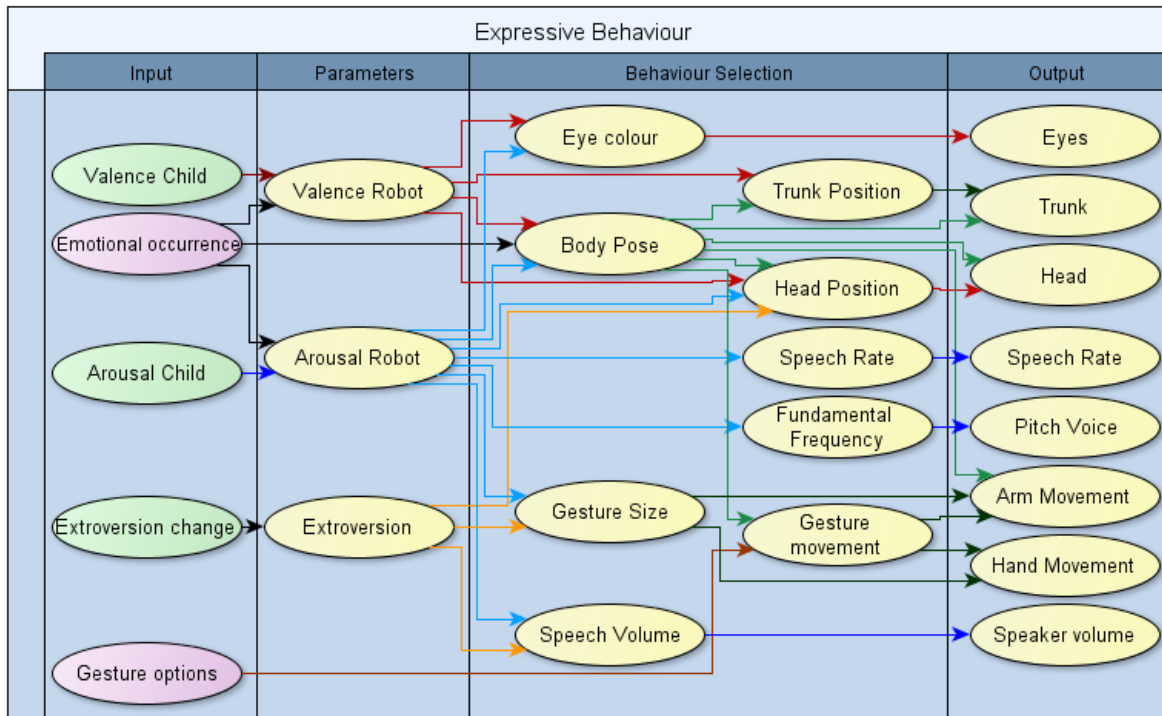


FIGURE 4.1: Visual representation of the information flow in the model

action sequences established in these use-cases as well as the corresponding requirements can be found in Appendix 1. The next stage in the sCE process would be the build stage, which will be presented in the next chapter. The aim in this chapter is to develop a model which can be entered into the build stage without the need of further specification. Since the model presented through use-cases alone does not meet this demand the details of this model need to be further specified. The first step in this specification is to identify the specific stages in the information flow based on the use-cases. The first stage identified is the input stage, in which what information enters the system is specified. The second stage is to adjust internal parameters based on the input. The third stage is the behaviour selection, where several modules work together to select behaviours based on the input and internal parameters. The final stage is to translate these behaviours into output. In the following sections I will discuss which processes take place in the different stages. A visual representation of this model can be found in Figure 4.1, a full-size version of this diagram in Appendix 2.

4.2.1 Input

Based on the specifications from section 3.4 we can identify the input for the system. The input will consist of those data which are used to set internal parameters and select appropriate behaviour. Two different kinds of input depending on their source can be selected. The first is the input from the child, those data which depend on the speech and emotional values from the child's state. In the diagram this kind is represented by the green modules. For the child's emotional state two kinds of input are needed, namely his valence and arousal values. The

second category of input is information about the task state, represented in purple. The other category is the dialogue selected, which is needed for the emotional and the gestural system and which is represented in blue. The emotional system receives information about the emotional relevance of the situation and the gestural system information about the gestures which fit with the dialogue.

4.2.2 Internal Parameters

Three internal parameters will be used which influence the behaviour selection process. The first parameter is extroversion, this determines how extrovert or introvert the robot is. This value is based on the extroversion of the child and is in theory not changeable. In practise it might, however, turn out that the extroversion of the child was not correctly tested. For this reason the extroversion parameter might have to be changed if it turns out to be unsuitable. The two other internal parameters represent the arousal and the valence of the robot. Arousal of the robot is dependent on the arousal of the child and the current task state. The arousal of the child influences that of the robot in that the robot's arousal will follow the arousal of the child. If the arousal of the child is higher than that of the robot, the robot's arousal will become higher and the other way around. There is a limitation though, since we don't want the arousal of the child to become too high or too low. For this reason there is an upper and a lower boundary of the child's arousal where the arousal of the robot will no longer follow, but instead move towards a set value. This way, if the child becomes too excited the robot will display calmer behaviour, if the child is bored or sleepy the robot will become more aroused. The current task state influences the arousal of the robot in that if there is an occurrence correlated with high arousal, the arousal of the robot will rise and the other way around. The valence of the robot works in almost the same way as its arousal. The first difference is that it (of course) follows the valence of the child instead of its arousal. The second difference is that there is no upper bound for the valence of the child where the robot will lower its valence. Also, the exact boundary for a valence which is too low is different from the boundary for arousal as a low valence is a greater problem than a low arousal.

4.2.3 Behaviour Selection

The most crucial phase is the behaviour selection phase, in which behaviours are selected based on the input and the internal parameters. This phase consists of several components, some which interact with each-other. Based on these interactions we can identify the module which regulates prosodic voice change and the module which regulates movement. When it comes to voice regulation, the module deciding on speech volume, fundamental frequency and speech rate are activated. Speech volume depends on extroversion and arousal of the robot. The

higher these values, the higher the volume will be. There is an upper and lower bound so that the speech will never be annoyingly loud or too soft. Fundamental frequency and speech rate both depend only on arousal of the robot, the higher the arousal the higher their values. These modules also have upper and lower bounds to ensure understandability. Eye colour is determined by arousal, high arousal will lead to warmer colours, low arousal to cooler colours. Eye colour will not be related to valence, since there is evidence that lighter colours are perceived more positively than darker colours. Since our robot is white, however, no light will result in a white eye, which is the opposite of what we would want (Boyatzis & Varghese, 1994; Kaya & Epps, 2004). The final group of modules regulate the movement selection. This group is the most complicated since it determines emotional body language and gesturing. The first module which is selected is the one which selects a whole-body pose. This module receives input from the arousal and valence of the robot and from the current task state. If there has just been an emotional relevant occurrence this module will be activated and will select a full-body pose based on the arousal and valence of the robot. If this pose is selected it is executed and until that time no other movement will be executed. If no body pose is selected, the modules for head position, trunk position and gesture movement will be activated. The module for head position receives information from the extroversion parameter, the arousal of the robot, the valence of the robot, and the dialogue. When the robot looks at the child, the position of the head in terms of up/down will be determined by the arousal and extroversion of the robot. If extroversion is low, the robot will look down more often. When the head is up, the robot's arousal and valence determine how high it is. The higher the arousal, the higher the head. There are, however, upper and lower bounds to the influence of arousal to ensure that the robot still appears to be looking at the child. The trunk-position module is also activated when a whole body-pose is not chosen and gets additional information about the valence of the robot. When the valence is high, the trunk will be stretched, when it is low the trunk will be bowed. There are two modules for gestures, namely gesture movement and gesture size. Gesture movement is activated when no whole body pose is chosen and determines which kind of movement will be executed based on the dialogue input. Next, the size of the chosen gesture will be determined based on the arousal of the robot and its extroversion level. The higher arousal and extroversion, the larger the movements will be. If extroversion is very low, this will mean that some gestures are so small that they are barely visible.

4.2.4 Output

The final stage of the model is the output. The first type of output are the values for speaker volume, pitch and speech rate. These are the variables of the Nao robot which represent speech volume, fundamental frequency and speech rate respectively. The second type of output is in the form of movement. If a body pose is selected, this body pose is executed and the only type of

movement while it lasts. If this pose is not chosen, the module for head position will determine the movement of the head, the trunk position module will determine the trunk movement and the gesture modules determine the hand and arm movements.

4.3 Dependencies

In the previous section the information flow is described, as well as which modules are dependent on which values. This description paints the global picture of the model, but to come to an architecture which is directly implementable one final step is necessary. This final steps consists of formalising the scales for all variables which need a numerical value and to specify the formulas which describe the dependencies between these variables.

4.3.1 Valence Robot

The first step in specifying how the valence of the robot is determined is to set a scale. For this model the choice is to set the scale from -1 to 1. The reason for this choice is that the software Facereader¹ uses this scale when calculating valence based on a person's facial expressions. Facereader is currently one of the state-of-the-art programs for emotion recognition and the Nao robots used within ALIZ-e will eventually have to use programs such as this to determine the emotional state of the child. Since the valence of the robot is largely dependent on the valence of the child it is useful to adopt the same scale for both values. The next step is to decide in which way the valence of the robot is dependent on the current task state and the valence of the child. To specify these dependencies, the following formulas can be drawn up for the different situations:

1. When the valence of the child is below X

If the valence of the robot is lower than 0:

$$\mathbf{VRnew} = \mathbf{VRold} + ((0 - \mathbf{VRold}) / 2) \text{ All values are rounded up to 1 decimal.}$$

If the valence of the robot is higher than 0:

$$\mathbf{VRnew} = \mathbf{VRold} - ((\mathbf{VRold} - 0) / 2) \text{ All values are rounded down to 1 decimal.}$$

Where VR_{new} is the new valence of the robot, VR_{old} the old valence of the robot .

¹<http://www.noldus.com/human-behavior-research/products/facereader>

2. When the valence of the child is not below X

If the valence of the robot is lower than that of the child:

$$\mathbf{VR_{new}} = \mathbf{VR_{old}} + ((\mathbf{V_{child}} - \mathbf{VR_{old}}) / 2) \text{ All values are rounded up to 1 decimal.}$$

If the valence of the robot is higher than that of the child:

$$\mathbf{VR_{new}} = \mathbf{VR_{old}} - ((\mathbf{VR_{old}} - \mathbf{V_{child}}) / 2) \text{ All values are rounded down to 1 decimal.}$$

Where VR_{new} is the new valence of the robot, VR_{old} the old valence of the robot and V_{child} the valence of the child.

3. When an emotionally relevant occurrence takes place, with valence Y:

If the valence of the robot is lower than Y:

$$\mathbf{VR_{new}} = \mathbf{VR_{old}} + ((\mathbf{Y} - \mathbf{VR_{old}}) / 2) \text{ All values are rounded up to 1 decimal.}$$

If the valence of the robot is higher than

$$\mathbf{Y: VR_{new}} = \mathbf{VR_{old}} - ((\mathbf{VR_{old}} - \mathbf{Y}) / 2) \text{ All values are rounded down to 1 decimal.}$$

Where VR_{new} is the new valence of the robot, VR_{old} the old valence of the robot and Y the valence of the emotion belonging to this particular occurrence.

The X value mentioned is the value below which the valence of the child becomes too low and where we therefore wish the robot to not mimic the child anymore. If this is the case the valence of the robot will approach the neutral value of 0. It is higher than the valence of the child, but not so high that the child may think the robot is glad because he's sad. When emotion recognition software is used, this value will have to be personalized for every child since different facial features can result in higher or lower valence values. Some children may simply appear to feel more negative than others. Also, a timer might have to be built in to prevent errors. The recognition software as it is today shows fluctuating interpretations and a child which just looks sad for a split second might not actually be sad at all. For this reason the mean emotion of a set period might be used instead of every millisecond of the analysis. When a wizard is used to interpret the emotion of the child, however, we would not need such a timer. It is also possible to set X to a specific value as the wizard can decide if the valence is too low. For this model, the chosen value of X for this situation is -0.75, which is based on the observation that current valence values from Facereader for very sad faces lie around -0.85. We do not wish the robot to start reacting when the child is already very sad, hence the choice for -0.75. The Y value is different for the different occurrences and depends on the emotion associated with that

occurrence. If the emotion is happiness, Y will be 1 as this is the value associated with a purely happy face in Facereader. If the emotion is sadness, Y will be -0.85, if it is anger, Y will be -0.8, if it is fear Y will be -0.5. These values are based on observations of Facereader ratings of sad, angry and scared faces. When an emotional relevant occurrence takes place the robot will still also mimic the child, but the occurrence will have more influence. This is modeled by simply adapting to the valence of the child first and then to the emotional occurrence.

4.3.2 Arousal Robot

The scale for the arousal of the robot will be set from -1 to 1. This choice is made because the same scale is used for the valence of the robot and it is practical to keep the scope for both values the same. The formulas for the arousal of the robot are very similar to those for valence, with the exceptions that the arousal of the child can also be too high, the X value is different and that in the case of a too high/low arousal the robots arousal will not go toward the same value of 0. The changes to the arousal of the robot are expressed by the following formulas:

1. When the arousal of the child is below X

If the arousal of the robot is lower than -0.7:

$$\mathbf{AR_{new}} = \mathbf{AR_{old}} + ((-0.7 - \mathbf{AR_{old}}) / 2) \text{ All values are rounded up to 1 decimal.}$$

If the arousal of the robot is higher than -0.7:

$$\mathbf{AR_{new}} = \mathbf{AR_{old}} - ((\mathbf{AR_{old}} + 0.7) / 2) \text{ All values are rounded down to 1 decimal.}$$

Where AR_{new} is the new arousal of the robot, AR_{old} the old arousal of the robot .

2. When the arousal of the child is above Z

If the arousal of the robot is lower than -0.4:

$$\mathbf{AR_{new}} = \mathbf{AR_{old}} + ((-0.4 - \mathbf{AR_{old}}) / 2) \text{ All values are rounded up to 1 decimal.}$$

If the arousal of the robot is higher than -0.4:

$$\mathbf{AR_{new}} = \mathbf{AR_{old}} - ((\mathbf{AR_{old}} - -0.4) / 2) \text{ All values are rounded down to 1 decimal.}$$

Where AR_{new} is the new arousal of the robot, AR_{old} the old arousal of the robot .

3. When the arousal of the child is not below X or above Z

If the arousal of the robot is lower than that of the child:

$$\mathbf{AR}_{\text{new}} = \mathbf{AR}_{\text{old}} + ((\mathbf{A}_{\text{child}} - \mathbf{AR}_{\text{old}}) / 2) \text{ All values are rounded up to 1 decimal.}$$

If the arousal of the robot is higher than that of the child:

$$\mathbf{AR}_{\text{new}} = \mathbf{AR}_{\text{old}} - ((\mathbf{AR}_{\text{old}} - \mathbf{A}_{\text{child}}) / 2) \text{ All values are rounded down to 1 decimal.}$$

Where \mathbf{AR}_{new} is the new arousal of the robot, \mathbf{AR}_{old} the old arousal of the robot and $\mathbf{A}_{\text{child}}$ the arousal of the child.

4. When an emotionally relevant occurrence takes place, with arousal Y:

If the arousal of the robot is lower than Y:

$$\mathbf{AR}_{\text{new}} = \mathbf{AR}_{\text{old}} + ((\mathbf{Y} - \mathbf{AR}_{\text{old}}) / 2) \text{ All values are rounded up to 1 decimal.}$$

If the arousal of the robot is higher than

$$\mathbf{Y}: \mathbf{AR}_{\text{new}} = \mathbf{AR}_{\text{old}} - ((\mathbf{AR}_{\text{old}} - \mathbf{Y}) / 2) \text{ All values are rounded down to 1 decimal.}$$

Where \mathbf{AR}_{new} is the new arousal of the robot, \mathbf{AR}_{old} the old arousal of the robot and \mathbf{Y} the arousal of the emotion belonging to this particular occurrence.

When the arousal of the child is too low, the arousal of the robot will approach -0.7. This value is below 0 since a low arousal is far less negative than a negative valence. If the arousal of the child is too high the robot will also approach a low arousal, but this will be value of -0.5 because otherwise the arousal of the robot might be too distant from the arousal of the child. For the values of X and Z in these formulas the same thing holds as for the X value in the valence formulas, namely that they will have to be personalised for each child when using emotion-recognition software. Facereader does not calculate arousal of facial emotions and we can therefore not simply look at arousal values for specific emotions to determine when arousal is too high or too low. It is possible, however, to calculate a measure for arousal in a similar way as to how valence is calculated. This method is to take the highest value for an emotion with a high arousal (happiness, anger, surprise and fear) and subtract the highest value for an emotion with a low arousal (neutral, sadness). Based on the values for arousal calculated in this way, the X value will be set to -0.95 and the Z value to 0.9 when the model works with a wizard interpreting the emotions of the child. The Y value is different for the different occurrences and depends on the emotion associated with that occurrence. If the emotion is happiness, Y will be 0.9, based on arousal values for happy faces. If the emotion is sadness Y will be -0.7, if it

is anger Y will be 0.8 and if it is fear Y will be 0.8. These values are based on observations of Facereader ratings of sad, angry and scared faces. When an emotional relevant occurrence takes place the robot will still also mimic the child, but the occurrence will have more influence. This is modeled by simply adapting to the arousal of the child first and then to the emotional occurrence.

4.3.3 Prosody voice

Three prosodic characteristics changeable in the Nao robot are speaker volume, speech rate and pitch, which correspond to speech volume, speech rate and fundamental frequency. In order to keep the Nao understandable it is important to set the right scales over which these characteristics can fluctuate. The speaker volume will be between 45 and 65 %, the speech rate between 75 and 115 and the pitch between 1 and 1.3. Note that these values might have to be changed depending on the environment. If a room is very noisy, for instance, speaker volume might need to be higher. All values are linear scales and depend on arousal, speech volume also on extroversion. The formulas for these values are:

Speaker volume:

$$SV = 45 + (((A + 1) / 2) * 20)$$

Where SV is speaker volume and A arousal

Speech rate:

$$SR = 75 + (((A + 1) / 2) * 40)$$

Where SR is speech rate and A arousal

Voice shaping:

$$VS = 1 + (((A + 1) / 2) * 0.3)$$

Where VS is voice shaping and A arousal

4.3.4 Eye colour

The eyes of the Nao robot contain full colour RGB LEDs, so it is possible to turn them into any RGB colour. The colour of the eyes is determined by the arousal of the robot. The higher the arousal, the warmer the colour. In order to specify the exact correlation between arousal and the RGB code it is necessary to first know how to express warmth in colour codes. An RGB value consists of three separate hexadecimal numbers, the first representing red, the second green and

the third blue. When considering these values in decimals, each colour has a scale from 0 to 255. For arousal, the highest arousal value will have a purely red colour, the lowest a purely blue one. The scale from red starts with the value R255 G0 B0 and moves through R255 G255 B0, R0 G255 B0 and R0 G255 B255 to R0 G0 B255. Since the scale for arousal is rounded off to one decimal, we can convert this scale into a scale from 0 to 100% with 5% steps. This gives us 21 different arousal percentages and it is possible to give every possible percentage its own colour taking in account the colour scale just mentioned. This list can be found in Appendix 3.

4.3.5 Trunk position

The Nao robot cannot move its shoulders, so trunk position is defined only by how much the robot bows forwards or leans backward. As it is important that the robot keeps its balance while adapting its trunk position, there is a limited amount of movement involved. The movement is implemented in the hips with the HipPitch joints. The position in which the robot leans back most corresponds to a HipPitch of 14, the position in which the robot bends forwards most with a HipPitch of 8. The formula specifying the relation between valence and HipPitch is the following:

$$\mathbf{HP} = \mathbf{6} + (((\mathbf{V}+1)/2)*\mathbf{6})$$

Where HP is the HipPitch value and V is the valence of the robot

4.3.6 Head position

Head position of the robot depends on emotional input. The head position of the robot will be determined by its arousal, valence and extroversion. Extroversion determines how often the robot looks down. The robot will look down between 0 and 10 times every 5 minutes. Every 30 seconds the robot will decide if it looks down with a chance of X, where X is defined by the following formula:

$$\mathbf{X} = (\mathbf{100} - \mathbf{E}) / \mathbf{100}$$

Where X is the chance of the robot looking down and E is the extroversion of the robot in percentages

Looking down will mean that the robot will lower its head by adding 15 points to its HeadPitch. It will go towards this position in 15 frames, hold it for 20 frames and go back to its original position in 15 frames. The robot thus looks down the same amount no matter what the start position of the head was. The start position of the head depends on the arousal and valence of the robot. A neutral position means that the HeadPitch is 3, which corresponds to an arousal

	Iconic	Deictic	Metaphoric	Beat
Extranarrative	0.1726	0.1151	0.1371	0.5752
Narrative	0.3694	0.3911	0.0222	0.2173

TABLE 4.1: The chance of each kind of gesture depending on the type of clause.

and valence of 0. The lowest head position corresponds to a HeadPitch of 16, the highest to -10. These values are chosen in such a way that there is a clear difference in head positions while the robot still appears to be looking ahead. Again, these values might have to be adapted if the child is sitting higher or lower than the robot. The formula for the HeadPitch of the robot is the following:

$$HP = -10 + (((AP + VP) / 200) * 26)$$

Where *HP* is the HeadPitch value, *AP* is the arousal of the robot in % ($AP = ((A+1)/2)*100$) and *VP* the valence of the robot in % ($VP = ((V+1)/2)*100$).

4.3.7 Gesture movement

The gesture movement module does not generate the movements based on a textual analysis in this model. This means that the movements are either included in the input from the dialogue or the options available for each piece of dialogue are known. To avoid a robot gesturing in exact the same way for every sentence it is a good idea, though, to have several gesture movements for every piece of dialogue. The gesture movement module has the task to decide which of the possible gestures to execute. There is more to this process than randomly picking a movement, though. Not all different types of gestures occur equally often in human conversation and this should be taken into account. Research by McNeill (1992) has shown that iconic and beat gestures are the most common, each making up approximately one third of gestures, about 5.5% of gestures is metaphoric, 3.5% deictic and 24% of gestures cannot be classified in one of the four categories. Moreover, the type of clause influences which gesture occurs, iconic and deictic gestures are used most for narrative clauses (+- 85% vs. 90% resp.) and metaphoric most in extranarrative clauses (+- 70%). Beat gestures occur equally often in both kinds of clauses. If we translate these percentages to make them add up to 100 we can specify how much chance every kind of gesture has to be chosen based on the clause of the dialogue. The exact chances can be found in Table 4.1.

4.3.8 Gesture size

Gesture size depends on the arousal and the extroversion of the robot, both of which have an equally large influence on gesture size. If both values were to be minimal, no gesturing would

occur. Gesture size is expressed in percentages, 100% being the gesture with the largest trajectory. The formula for gesture size is:

$$GS = (((A + 1) / 2) * 100) + E) / 2$$

Where *GS* is the gesture size in percentages, *A* is the arousal of the robot and *E* the extroversion

4.4 Timing

Now that the information flow, the different modules and the dependencies between values are clear, only one question remains with respect to timing. Values are updated in the model, in theory this could be done several times a minute, but is that also what we want? The formulas for mimicking the valence and arousal of the child are designed in such a way that the robot will go halfway towards the valence and arousal. The reason for this is that we do not want the robot to exactly copy the child's emotion, only be influenced by them. If this formula is executed several times per second, however, this is exactly what would happen. The robot would update its valence and arousal several times in the direction of that of the child until it is the same, all within a couple of seconds. This will give the appearance of the robot copying the child's emotion. For this reason, a timer will have to be built in. This timer will start to run whenever the arousal or valence of the child changes. If this happens, the corresponding formulas will adapt the arousal and valence of the robot. After this change, the timer will have to have reached a certain value before the arousal and valence can be updated again. The exact value of this timer will have to be established through trial-and-error, but I suggest 30 seconds as a starting point. This is low enough that the robot will still display changing behaviour but not so low that the robot will copy the child. Through this timer, the robot will adapt its valence and arousal to the child, but gradually. Eye colour, trunk position, the voice characteristics and gesture size all depend on the values of arousal and/or valence. This means that they will follow this slower change in these values. The gesture movement module will only be activated when a dialogue with corresponding movements is given as input.

4.5 Use-case

To illustrate how the model will operate in a situation with a task and a child, a part of a use-case will be presented in this section. This specific use-case is based on a quiz-scenario, the robot and the child play a quiz in which they can ask each other questions. It presents an approximation of how the different behaviours would change during an interaction with a child. The reasons that this is only an approximation are, firstly that the reaction of the child

	22 Child speaking	23 Model	24 Child speaking	25 Model	26 Child speaking	27 Model
Arousal Child		-0.2		0		-0.1
Valence Child		-0.75		-0.4		-0.2
Dialogue robot		Antwoord 1?		Hoera! Ik sta nog steeds voor. Het is alweer tijd om te stoppen. Vond je het leuk?		Wat jammer, ik vond het wel leuk.
Task state		Neutral		Happy, A0,9 V1		Sad, A-0,7 V-0,85
Speech child	Wat is gezonder, olie of boter? Antwoord 1: Olie, dat zijn gezonde vetten, 2: Boter, want dat is gemaakt van melk		dat is goed		nee	
Extroversion		40		40		40
Arousal Robot		-0.2		0.4		-0.3
Valence Robot		-0.1		0.4		-0.4
Eye colour						
Speaker Volume		53		59		52
Pitch		1.12		1.21		1.11
Speech Rate		93		99		92
Head position		0.4		Happy		Sad
Trunk position		10.5		Happy		Sad
Gesture Move.		Iconic		Happy		Sad
Gesture Size		40		Happy		Sad

FIGURE 4.2: Visual representation of the information flow in the model

on the situation and the robot had to be guessed. The second reason is that many values will be updated independent of the dialogue state. Since it is impossible to review every second of an interaction, this use-case will presume that values are only updated right before the robot speaks. Only a short part of the use-case will be presented here due to the large number of parameters updated in each step. The values of each parameter for each step can be found in Figure 4.2. The full use-case for the short quiz can be found in Appendix 4, the numbers at the top of Figure 4.2 are references to the corresponding line numbers in Appendix 4.

Figure 4.2 presents the end of a quiz game between robot and child. The first moment (Nr. 22) is the child asking the robot a question. The next moment starts with the input. The child had just answered the second question in a row wrongly and is therefore feeling quite sad. The robot will give an answer to the child and the task state is neutral. Based on the input, the arousal of the robot follows that of the child. The valence of the robot goes towards 1 because

the valence of the child is very low. The extroversion of the robot stays 40. In the next step the eye colour of the robot is generated (green) and speech volume, fundamental frequency, speech rate, head position, trunk position and gesture size are calculated with the formulas presented in the previous section. An iconic gesture is chosen to accompany the speech. In the next moment the robot has uttered its dialogue and the child is replying. In moment 25 the input shows that the arousal and the valence of the child have risen again. Since the robot has just answered a question right, information from the task state is given about the corresponding arousal and valence values. The arousal of the robot is adapted first to the child and next to the happy occurrence. Next, its eye colour is generated (light yellow/green) and speech volume, fundamental frequency and speech rate are calculated. Since a happy occurrence has just taken place the robot will not generate a head position, trunk position or gesture but instead display a happy pose. In the next moment the child gives an answer to the question of the robot. In the final moment, the arousal and valence of the child have again dropped a bit. Since the child has just responded negatively to the question if he had liked the game, the robot will display a sad pose. Based on the input from the child the arousal and valence values are adapted. Based on these values and extroversion the eye colour is calculated (dark green) and the speech volume, fundamental frequency and speech rate calculated. The movement modules will not calculate anything since a sad pose will be displayed.

4.6 Conclusion

In this chapter I have presented the outline of a model for expressive behaviour for the Nao robot. I have first specified which stages in the information flow exist and what the role is of the different modules. Next, I have specified the model by formalizing the dependencies between values and modules. Finally, I have given a short example of how this model would work in a scenario.

Chapter 5

Implementation

5.1 Introduction

The model presented in the previous chapter was implemented so it could be used in an experiment and its effects could be observed. In this chapter I will present the way in which the model was implemented.

5.2 GOAL

The programming language used within TNO for the reasoning programs of the robot is GOAL. This is therefore the language which was used to implement the model presented in the previous chapter. GOAL is a prolog-based BDI agent language developed by Koen Hindriks¹. A GOAL program consists of a knowledgebase with static facts, a belief base with changeable beliefs, a goal base with changeable goals, an action base which specifies actions and modules with rules for which action should be performed based on percepts of the environment, beliefs and goals. GOAL is a language particularly suited for reasoning about which behaviour to display in which situation. This focus on behaviour selection instead of behaviour generation means that the individual behaviours are not necessarily fully specified in the program, but can simply be represented by constants. A specific iconic gesture for instance, is represented by *i1*, *i2* etc. The Goal program will give this constant as output and a behaviour with the same name is known to the Robot. The GOAL program thus decides which behaviours are appropriate in which situation.

¹GOAL Website: <https://mmi.tudelft.nl/trac/goal/wiki>

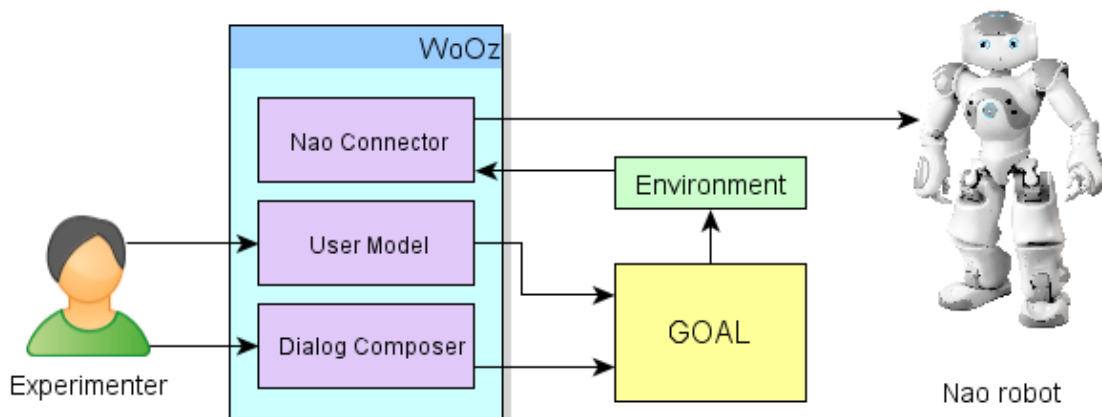


FIGURE 5.1: The different components in the system and how they communicate

5.3 Communication

In order for a GOAL program to be able to reason about events outside its scope, it needs to be able to receive input from its environment and to send output back. This is particularly relevant to this project as we wish for the problem to reason about actions which should be performed real-time. Within the Aliz-e project, different technical components can be distinguished. The first is, of course, the robot itself. This robot is not yet completely autonomous, so to make experimenting with those components which are finished possible, a simulation of the missing parts is needed. For this reason, the second component used is a so-called Wizard of Oz (WoOz), which is operated by an experimenter. The third component is the GOAL program, which runs automatically. Because not all input which our model needs can be automatically generated at this time, the WoOz component will provide the input to GOAL. The WoOz thus serves as a communicator between the robot and the GOAL program. In this section, I will specify the different components at work in this communication. In Figure 5.1 a schematic representation of the different components can be found.

5.3.1 Wizard of Oz

The Wizard of Oz interface is an interface which allows an experimenter to send commands to the robot in an intuitive way. The WoOz also serves as a communicator between GOAL and the robot. All percepts are sent to GOAL through the WoOz and all actions from GOAL are first processed by the WoOz, which then sends them to the robot. The WoOz environment used by TNO within the Aliz-e project was developed by Bert Bierman ², who also integrated

²Produxi <http://www.produxi.nl/>

all changes necessary for the WoOz to be able to connect with GOAL. The WoOz consists of several modules, of which the most important for this project are the User Model and the Dialog Composer. The User Model represents all information about the user, in this case the child. Through the User Model, the experimenter can send commands to GOAL about the emotional state of the child. The User Model also has fields with which to adapt the range for head position, trunk position, speech volume, speech rate and fundamental frequency. These commands are included because the appropriate ranges are situation dependent. Head position, for instance, depends on the height of the user in relation to the robot. Although these ranges could simply be altered in the GOAL program itself, I have chosen to include them in the WoOz as well. The reason for this is that the model presented in this thesis is meant for use within the ALIZ-e project. This means that in the future, people without any knowledge of GOAL might work with this model. As the WoOz interface is much more intuitive than GOAL code, it is preferred if all changes which might have to be made to the model can be communicated through the WoOz. The second important component is the Dialog Composer. Through this component texts are selected and sent to the robot. With these texts, situation related commands and gestures are sent to GOAL when they are selected.

5.3.2 Environment

An *environment* for a GOAL program specifies the results of the actions performed. In the case of my program, it is the environment through which the GOAL program sends information back to the WoOz. An ALIZ-e environment already existed before I started my project ³, but due to specific demands it had to be adapted. A GOAL environment is a .jar file which consists of several class files and additional .jar files. One of these class files specifies the GOAL actions which can be sent back to the WoOz. My adaptations to the ALIZ-e environment consisted of specifying the names and number of arguments for the actions new in my GOAL program in the corresponding .java file. This file was compiled and the .class file corresponding to it was added to the original environment .jar file.

5.4 Program

A GOAL program consists of several components which work together to decide on which actions to perform. Because of the size and complexity of the final program, it is not possible to review every piece of code in detail in this chapter. Instead, I will give an overview of the functions implemented in each section. A full transcription of the GOAL code with clarifications is given in Appendix 5.

³Original environment was created by Wouter Pasman

5.4.1 Knowledge Base

The knowledge base of a GOAL program is static and therefore used to specify rules and facts which cannot change while the program is running. The code in the knowledge base is only used in the program when it is called from the main module, when variables will be instantiated. Several categories of rules and facts can be distinguished.

- A formula which generates a random number between 0 and 100. This random number can later be used to calculate chances. For instance, if we wish for a 50% chance for A, we could say that A happened when the random number generated was below 50.
- Formulas which deal with time. For instance, after a certain amount of time has passed with no change in the emotions of the child, we wish for the robot to again adapt its valence and arousal. We also wish for the robot to decide if it has to look down every 30 seconds.
- Rules which define when the arousal of the child is too high or too low and when the valence is too low.
- Two rules which calculate the arousal and valence of the robot in percentages. For some formulas, it is easier to use a scale from 0 to 100 than from -1 to 1.
- The formulas which define the behaviour of the robot. As seen in chapter 4, formulas can be given for the values of speech volume, speech rate, fundamental frequency, head position, trunk position and eye colour. These formulas are all defined in the knowledge base.
- Formulas which state how the arousal and the valence of the robot need to be adapted. When the arousal of the robot has to be adapted towards a higher value, the formula which specifies this change is different from when it has to be adapted towards a lower value. For this reason, we need to specify when to use which formula.
- Rules for choosing a gesture. Not all gestures have an equal chance to be chosen by the system. Beat gestures are, for instance, more likely to occur than deictic gestures and the initial chances are saved in the belief base. When one gesture cannot occur, however, we would still like the total chance of any other type to be 1. The rules in the knowledge base specify how to calculate the updated chances, which gesture is eventually chosen and if multiple of this type exist, which ID number it has.

5.4.2 Belief Base

The belief base represents the beliefs of the robot. Beliefs can change while the program is running, they can be removed and new ones can be added. In this section, I will only discuss those beliefs which are in place when the program is started. Note that the belief base will constantly change when the program is running. All rules for when beliefs are added or removed are found in the Main program or the Event module. I will list the types of beliefs here.

- We have a start predicate which is removed when the program begins, but which is needed so the program knows to initialize the start time.
- Beliefs which deal with the internal state of the robot, these are beliefs about its extroversion, arousal and valence.
- Beliefs about the child, the robot has a belief about the arousal and about the valence of the child.
- Beliefs about the current behaviour displayed. The GOAL program knows the last values of speech volume, speech rate, fundamental frequency, head position, trunk position and eye colour which have been sent to the robot. These values are used to check if they need to be updated. If the values do not correspond to the values calculated by the knowledge base (based on the current arousal and valence of the robot), the robot knows they need to be revised. When the program starts, the values will be set to 0, so the robot will always update its behaviour immediately.
- Beliefs about the maximum and minimum values for head position, trunk position and voice characteristics. These are in the belief base so they can be changed from the WoOz. Head position, for instance, might be dependent on how high the robot is placed on a specific table. This is particularly useful when the experimenter operating the WoOz is not familiar with GOAL.

5.4.3 Actions

The actions in a GOAL program generate its output, the actions communicate changes to the WoOz through the environment. One action for instance gives a value for the speech volume of the robot. The environment conveys this value to the WoOz, which in turn communicates it to the robot, who changes its speaker volume to the value received. The actions in my program specify the different behaviours the robot needs to display. It gives values for speech volume, speech rate, fundamental frequency, trunk position, head position, eye colour, tells the robot to look down and which gesture to perform with which size.

5.4.4 Program Main

The main part of the program consists of rules for what should happen in which situation. They specify which actions to perform, but also when beliefs should be added or removed based on the current belief base. The rules in this section all have an *if X then Y* format. In the X, information from the belief and knowledge base is used to specify in which situation the rule applies. The Y specifies what should happen in that situation. This can include actions which are sent to the robot, but also internal functions such as removing or adding beliefs. Again, the different rules can be divided into groups based on what their function is.

- The first rules are time related. They state, based on the knowledge base rules about time, when arousal & valence need to be updated and when the robot looks down.
- The next category deals with updating the arousal and valence of the robot when these have not yet been adapted to the child or occurrence. Depending on whether or not the valence and arousal of the child are too high or low, arousal and valence are adapted in the belief base.
- The third type of rule deals with updating the behaviour of the robot. When the old value in the belief base does not match with the formula in the knowledge base (based on current arousal & valence), the robot updates its beliefs and performs the fitting action. This method is used for adapting speech volume, speech rate, fundamental frequency, trunk position, head position and eye colour.
- The final type are the rules which deal with gesture selection. These see which gesture has to be chosen based on the knowledge base, perform the fitting action and remove all known chances for gesture types from the belief base. (These will be re-instated once a new gesture has to be chosen)

5.4.5 Events

The last section of the program is the event section. This section is different from the Program Main in that in the Event section, each applicable rule will be executed every cycle, while in the Program Main only the first rule found is executed. For this reason, the event section is used to handle the input, or the *percepts* from the environments. These rules specify which beliefs should be added or removed based on the current state of the environment. Some rules which do not deal with percepts are added to this section because of its characteristic of executing every applicable rule. The additional rules are those rules which we would like to have performed quickly every time they are relevant. These are the rules which deal with gestures, because they

deal with inserting precepts into the belief base and the rule which removes the start predicate. The rules in the event module are:

- Percept rules which deal with the arousal and valence of the child. When a new percept enters the system, the old belief about the arousal or valence of the child is updated and a belief is added which indicates that the arousal and valence of the robot are not up to date any more.
- When the extroversion of the child changes, the robot will change its former extroversion value to that of the child.
- When an emotionally relevant occurrence takes place, the corresponding values for arousal and valence are added to the belief base so that the arousal and valence of the robot can be adapted to them.
- When a percept with gesture options occurs, the specific chances for each kind of gesture will be added to the belief base, based on what type of text they occur with. After this, a check is performed to see if any of the types of gestures is indicated by a 0, which means that no gesture of this type exists. If this is the case, the chance for this type of gesture in the belief base is changed to 0. Rules from the knowledge base make sure that the other chances will be evenly distributed in a later stage.
- Percept rules which deal with changes to the maximum or minimum values for head position, trunk position and voice values. When a new percept for one of these values occurs, the corresponding belief in the belief base is changed.
- The final rule will only be executed once, namely during the first cycle. When *start* is still in the belief base, the current time will be saved and the *start* predicate removed. This rule is thus used to initialize a time variable so that the time since the start of the program can be measured.

5.5 Conclusion

In this chapter I have presented the implementation of the model described in chapter 4. The work described corresponds with the *build stage* of the sCET methodology. The next step is the *evaluate stage*, which will be presented in chapters 6 and 7.

Chapter 6

Experiment Preparations

6.1 Introduction

The system presented in the previous chapters was developed with the aim of improving the interaction between robot and child. From the literature we expect it to have several effects, such as improving how much the child likes the robot and how much it feels understood by the robot. We cannot say with certainty, however, what the effect of the system on interactions actually is before this is tested. For this reason, an experiment was conducted aimed to test the effect of the adaptive expressive behaviour on the experience and (interaction) behaviour of children. Although the system of expressive behaviour is finished at this point, some further preparations were necessary to realise the actual experiment. This chapter reports these preparations.

6.2 Arousal and Valence

Although the model for expressive behaviour has been designed and implemented, it is still dependent on input about the emotional state of the child from other sources. In the experiment, an experimenter will fill this role using the WoOz. Because we do not want the performance of the system to depend on too subjective a measure, however, it is important to specify how valence and arousal are interpreted in as much detail as possible. The ideal is a script for the experimenter which interprets the behaviour as exact as possible, but we should take into account that this interpretation needs to be done real-time. This means that it should also be simple enough that an experimenter can use it very quickly.

Behaviour	Is linked to	Relation
Speech volume	Arousal	Louder is higher
Speech rate	Arousal	Faster is higher
Fundamental frequency	Arousal	Higher is higher
Gesture size	Arousal	Larger is higher
Head position	Arousal & Valence	Higher is higher
Trunk position	Valence	Straighter is higher

TABLE 6.1: Behaviours and their relation with Arousal and/or Valence

Behaviour	Is linked to	Relation
Smiling	Valence	More is higher
Laughter	Valence	More is higher
Frowning	Valence	More is lower
Crying	Valence	More is lower
Bouncing	Arousal & Valence	More is higher
Shrugging	Arousal & Valence	More is lower

TABLE 6.2: Behaviours and their relation with Arousal and/or Valence

6.2.1 Behaviour characteristics

In order to arrive at a script for interpreting the arousal and valence of a child it is important to know which observable behavioural signs can give an indication of these values. Since we have already established which behaviours the robot should adopt depending on valence and arousal, this is a good starting point. The different behaviours with their dependencies are listed in Table 6.1.

When it comes to valence, not many values in Table 6.1 refer to it. This has to do mostly with the fact that valence shows most clearly in facial expressions, which the Nao robot cannot perform and are therefore not included. A child, on the other hand, will display some behaviours linked to emotion which the robot can't, including facial expressions. It is therefore important to also select some behaviours unavailable to the robot which are important in interpreting emotion. For this reason, Table 6.2 was drawn up to supplement Table 6.1.

Although this table clarifies some observable variables related to valence and arousal, we cannot simply state *when speech volume is X then arousal is X*. When interpreting a value like arousal we need to take into account all different behaviours which have a relation to arousal. Another issue is that people cannot measure a variable such as speech volume exactly. It is therefore impossible to provide a direct relation between speech volume and arousal when a person is interpreting the value of the speech volume. This means that although an automatic system might be able to calculate the arousal of a child based on exact values of voice characteristics, a system for an experimenter needs to be more intuitive. Such a system needs to take into account the fact that people cannot measure each variable mentioned in the tables exactly.

In order to specify how to interpret behaviour, the behaviours in Tables 6.1 and 6.2 can be divided into two groups. The first group consists of those behaviours which are continually observable, namely speech rate, speech volume, fundamental frequency, gesture size, head position and trunk position. The other group consists of the other behaviours, which can occur for a moment and then stop again. Because of the number of different behaviours which express valence and arousal it is nearly impossible to follow an exact script to interpret them. If a person has to interpret emotion following 12 different parameters at the same time, there is a big chance that factors are missed. For this reason I have chosen not to specify the exact relation between any behaviour and an arousal or valence setting. It is possible, however, to set some guidelines which should be followed. (Obviously, common sense should still be used in special cases such as happy tears)

- The first guideline is a list of the behaviours mentioned in Table 6.1 and 6.2. Aside from the obvious facial expressions to interpret emotional state, these behaviours should be kept in mind when observing the child.
 - When you notice a rise in speech rate, arousal has gone up.
 - When you notice a rise in speech volume, arousal has gone up.
 - When you notice a rise in fundamental frequency, arousal has gone up.
 - When you notice an increase in gesture size, arousal has gone up.
 - When the head is higher than average, arousal and valence are above 0.
 - When the trunk is in a more stretched position than average, valence is above 0.
 - When the child smiles, valence is above 0 and goes up.
 - When the child laughs, valence is above 0.4 and goes up.
 - When the child frowns, valence is below 0 and goes down.
 - When the child cries, valence is below -0.4 and goes down.
 - When the child starts bouncing, arousal is above 0 and goes up.
 - When the child starts shrugging, arousal is below 0 and goes down.
- All of the before mentioned behaviours can also have opposite effects, examples listed below. Notice that the absence of certain behaviours (such as smiles) is also significant!
 - When you notice a decline in speech rate, arousal has gone down.
 - When head is lower than average, arousal and valence are below 0.
 - When the child stops smiling, valence goes down.
 - When the child stops shrugging, arousal goes up.
 - Etc.

- Arousal and Valence work on a scale from -1 to 1. This means that if they are both -1, the child cannot be any sadder and if they're both 1 the child cannot be happier. This should be taken into account when approaching these values. If Valence is already 0.9 the child should appear very happy, otherwise something might be wrong with the interpretation.
- The value of 0 for both Arousal and Valence represent the middle between the lowest and highest possible values. This means that Arousal 0 and Valence 0 do not necessarily represent the default emotional state of the child! It can very well be that the child is constantly calm and content, which would be better represented by, for instance, an Arousal of -0.4 and Valence of 0.3.

6.2.2 Testing Objectivity

It not possible to develop a system for the interpretation of valence and arousal which is fully objective. It would, however, be good to have an indication of how subjective the interpretation of arousal and valence actually is. In order to test this, a small experiment was conducted. The participants watched a video recording of a child playing with the Nao robot. While watching, they had to keep track of the arousal and valence of the child. This test enables a comparison of how different people interpret the emotions of a child showing exactly the same expressive behaviour.

Two methods of keeping track were used. The first were two separate sliders, one for arousal and one for valence, as shown in Figure 6.1. The valence and arousal can be changed either by pressing the *Up* or *Down* buttons, or by dragging the slider. The buttons raise or lower the value by 0.1, when the slider is dragged the same holds. It is not possible to click somewhere on the scale to move the slider to that point. Arousal and valence can therefore be changed only with 0.1 per click. The other input method is a two-dimensional model with valence on the X axis and arousal on the Y axis. This model also has the possibility of showing where some specific emotions lie on the scales. The emotional points were placed according to research by Russell (J. Russell, 1980). In this study, the emotions *alarmed*, *tense*, *angry*, *afraid*, *annoyed*, *distressed*, *frustrated*, *miserable*, *sad*, *gloomy*, *depressed*, *bored*, *droopy*, *tired*, *sleepy*, *calm*, *relaxed*, *satisfied*, *at ease*, *content*, *serene*, *glad*, *pleased*, *happy*, *delighted*, *excited*, *astonished* and *aroused* were rated along the pleasure-displeasure and high-low arousal scales. This scaling resulted in a two-dimensional picture with the emotion words at their respective values. Because the scaling in the original picture was done on a circular scale, the image was slightly adapted to form a square. Because 28 emotion points would make the image fairly full, a selection of 12 words was made. For each quarter of the picture, the three most common words were selected. This resulted in the subset *afraid*, *angry*, *annoyed*, *miserable*, *sad*, *bored*, *sleepy*, *calm*, *content*, *glad*, *happy* and *excited*. The final visual representation is shown in Figure 6.2.

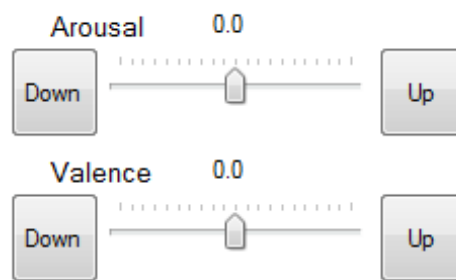


FIGURE 6.1: The first input method for Valence and Arousal

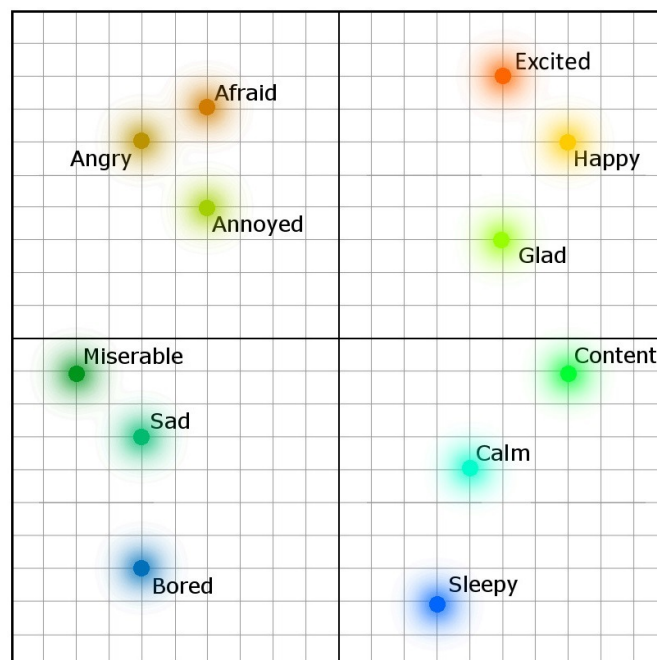


FIGURE 6.2: The second input method for Valence and Arousal

6.2.2.1 Method

Besides the author, six other students without prior knowledge of arousal and valence participated in the experiment. The author provided input two times for both input methods. Three of the other subjects used the first input method, the other three used the second input method. All subjects watched a video of a child interacting with the Nao robot in a quiz. The video lasted 4,10 minutes, during which the participants had to keep track of the emotional state of the child. The naïve subjects first got an introduction text, which explained the concepts of valence and arousal and their relation with behaviour. The full texts can be found in Appendix 6. Because the experimenter had much more knowledge starting this experiment, these results couldn't be compared with those of the other subjects. For this reason, the experimenter scored the video twice for each input method. The results from these four sessions should give us an

Subject	Mean Arousal	Lowest Arousal	Highest Arousal	Variance Arousal	Mean Valence	Lowest Valence	Highest Valence	Variance Valence
Exp 1.	0.2478	-0.3	0.7	0.0583	0.3110	-0.2	0.8	0.0822
Exp 2.	0.3912	-0.3	0.7	0.0399	0.3682	-0.3	0.7	0.0501
1	0.1810	-0.3	1	0.0844	0.1798	-0.4	0.9	0.1078
2	0.2116	-0.2	0.4	0.0246	0.1652	-0.3	0.5	0.0585
3	0.0652	-0.4	0.4	0.0383	0.1176	-0.4	0.9	0.0456

TABLE 6.3: Results - Input Method 1

Comparing	Mean difference 1 Arousal	Mean difference 2 Arousal	Mean difference 2 Valence	Mean difference 2 Valence
Exp 1 & Exp 2	0.1922	0.1717	0.2208	0.2297
1 & 2	0.2062	0.2053	0.2086	0.2105
1 & 3	0.2770	0.2778	0.2850	0.2855
2 & 3	0.1880	0.1492	0.2044	0.2010

TABLE 6.4: Results 2 - Input Method 1. Mean difference 1 = Mean of the differences between 1 & 2 at each second. Mean difference 2 = Mean of the differences between the difference to the mean of 1 and the difference to the mean of 2 each second.

indication of how much variation there is within one person. The other subjects all rated the video once, their ratings could be compared with one another. The results for input method 1 can be found in Tables 6.3, and 6.4, the results for input method 2 in Tables 6.5 and 6.6. In the comparison tables, two methods for comparing subjects were used. Taking, for example subject 1 and 2 we look at the difference between both arousal and valence values as given at each time interval. By calculating the mean of these differences, we get a measure for how much the input of subject 1 and 2 differed. These scores can be found under the header *Mean difference 1 Valence* and *Mean difference 1 Arousal*. One objection to this way of giving a difference-score is that one person might have always scored 0.2 lower than the other. This would give a mean difference of 0.2, while both subjects could have raised and lowered arousal and valence at exactly the same moments. Since the changes in emotion are more important than the exact value, another way of calculating difference is used. With this method, we look at the difference from the mean score given at each point. This way we can compare two subjects on how much higher or lower than average they scored the child at each moment. This difference can be found under the header *Mean Difference 2 Valence* and *Mean Difference 2 Arousal*. One disadvantage of the data in the tables is that it does not provide information about the changes in valence and arousal overtime for different subjects. For this reason, the graphs comparing the mean difference 1 between the trials with the experimenter for input method 2 are shown in Figures 6.3 and 6.4.

Subject	Mean Arousal	Lowest Arousal	Highest Arousal	Variance Arousal	Mean Valence	Lowest Valence	Highest Valence	Variance Valence
Exp 1.	0.3124	-0.3	0.7	0.0526	0.3196	-0.2	0.9	0.0444
Exp 2.	0.2112	-0.2	0.7	0.0228	0.2726	-0.2	0.7	0.0556
1	0.1726	-0.2	0.8	0.0277	0.0382	-0.5	0.5	0.0689
2	0.3802	-0.6	0.8	0.1197	0.3322	-0.4	0.6	0.0300
3	0.0912	-0.6	0.7	0.0978	0.1326	-0.5	0.6	0.0857

TABLE 6.5: Results - Input Method 1

Comparing	Mean difference 1 Arousal	Mean difference 2 Arousal	Mean difference 2 Valence	Mean difference 2 Valence
Exp 1 & Exp 2	0.2140	0.1994	0.1830	0.1851
4 & 5	0.3160	0.2682	0.3128	0.2148
4 & 6	0.2666	0.2591	0.3232	0.3172
5 & 6	0.4338	0.3659	0.2892	0.2457

TABLE 6.6: Results 2 - Input Method 1. Mean difference 1 = Mean of the differences between 1 & 2 at each second. Mean difference 2 = Mean of the differences between the difference to the mean of 1 and the difference to the mean of 2 each second.

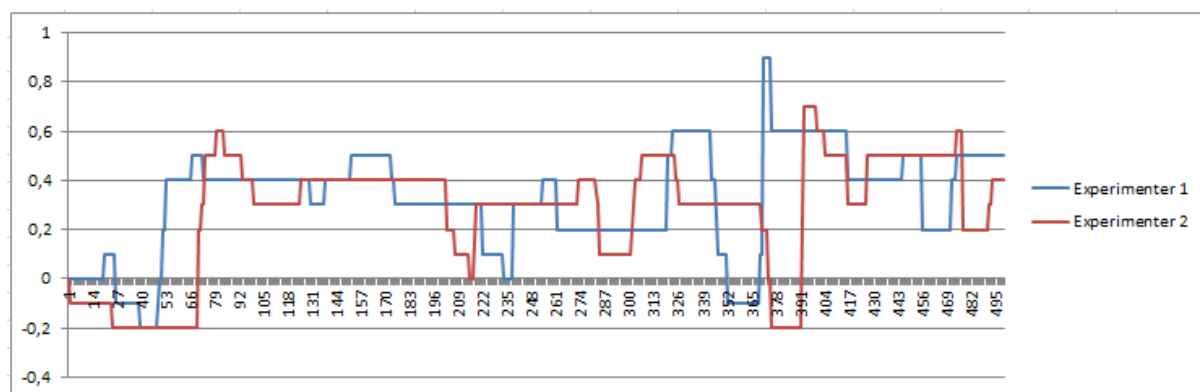


FIGURE 6.3: Valence

6.2.2.2 Discussion

These results show us several things, depending on which data we compare. First, we can compare the input methods on subjectivity. We do this by comparing the difference scores for both input methods. A comparison shows that there is no significant difference in subjectivity between the two methods, for either way of calculating difference score, for neither valence nor arousal. A second comparison is whether or not the within-subject differences are significantly smaller than the between-subject comparison. Since we have seen that the input method does not give a significant effect on the difference scores, we can use these for both methods. For both Valence and Arousal, and both difference measures, the difference within subject is lower than the differences between subjects. For none of the cases, however, a significant difference is found. A third comparison can be done based on the data per subject. This data shows

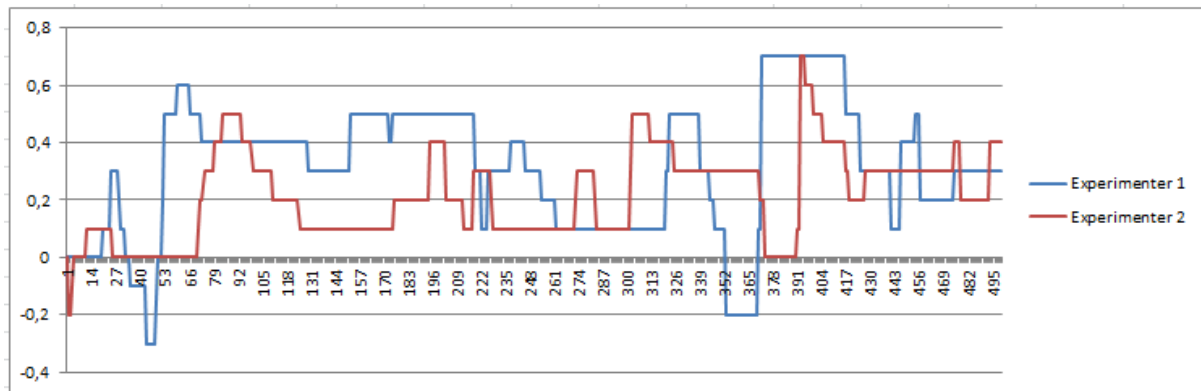


FIGURE 6.4: Arousal

that there are differences between subjects in means, variance and range for both arousal and valence. The final thing to be looked at are the two graphs displayed in Figures 6.3 and 6.4. These graphs show that although there are sometimes large differences per time point, the peaks and lows of the graph are mostly in the same place. This shows that a raise in valence or arousal was almost never interpreted as a drop and the other way around.

6.2.2.3 Conclusion

Concluding, we can state that this small experiment does not provide us with data which would justify choosing one of the input methods over the other. This could, of course, have something to do with the small amount of subjects which participated. I will take this as an indication, however, that there is no huge difference in subjectivity for both these methods. Since a choice did need to be made, however, other researchers from the ALIZ-e project were asked their preference. These discussions yielded the information that the second input method was preferred, because it was perceived as more intuitive. Based on these discussions the choice therefore fell on the second input method. Another conclusion we can draw is that it is not possible to let people interpret the emotions of a child and not get an input which will differ per person. The input will therefore always be subjective, even when provided by the same person. Exactly how subjective can best be seen from Figures 6.3 and 6.4. These graphs show us that although input can differ between sessions, the general trends will be the same. More importantly, a certain behaviour from the child will almost never be interpreted in two opposite ways.

6.3 Dialog

When the robot plays a quiz with a child, a certain amount of conversation is involved. While playing the quiz, the texts spoken by the robot are integrated in the quiz-interface of the WoOz. These are the questions asked, the answers given, comments on the answers of the child and comments on performance. Before and after the quiz, however, an introduction and a concluding conversation are needed. These conversations are transcribed in XML and will be sent to the robot through the dialog-interface of the WoOz. Two different introductory and two concluding dialogs are written, as children will play with both an expressive and a non-expressive robot. Both dialogs will be used by both robots in different sessions. Additional to these dialogs, each robot will tell a short joke half-way through the quiz. These jokes are also a part of the XML dialog transcriptions. All dialogs can be found in Appendix 7.

6.3.1 Gesturing

The gesture aspect of the system for expressive behaviour is closely linked to the text. As could be seen in Figure 4.1, the gesture system needs input about the possible gesture options. This input will be provided from the text of the Nao robot. The input takes the form of commands which are sent to GOAL the moment a certain dialog is selected. This can be both when questions are asked from the quiz-interface as when dialog options are selected in the dialog-interface. GOAL will then choose one of the gestures and send this gesture to the robot to be executed. A total of 27 gestures were created for this experiment, 8 beat gestures, 10 deictic gestures and 9 iconic gestures. Descriptions of the different gestures can be found in Appendix 8. Every piece of text spoken by the robot is transcribed with a subset of these gesture options, as well as a label *narrative* or *extranarrative* to assist with the gesture choice. The gestures were chosen to match the texts.

6.3.2 Emotional occurrences

During the interaction, emotional occurrences can take place. These are, for instance, the child answering positively to the question *did you like playing with me?*. At these moments, the robot will display the fitting emotion and its arousal and valence will be influenced by the emotion of the occurrence. The emotional occurrences will be indicated from the text. This is possible because the robot will always comment on the cause of such an occurrence. In the case of the example above, the robot could say *I'm glad, I liked playing too!*. At these points in the text two transcriptions will occur in the text. The first one will refer to the emotional body pose the robot is to display. The different body poses used were developed by Beck, Stevens, Bard,

and Canamero (2012) and Aldebaran ¹. The second transcription will be a command which will tell the GOAL program to adapt the arousal and valence of the robot appropriately. Whenever an emotional occurrence takes place with a piece of text, no gesture options will be given as an emotional pose cannot be combined with a gesture. This has to do with the fact that the arms of the robot follow a set trajectory for each emotional pose. When the robot is happy, for instance, it raises its arms. A gesture is a specific arm movement, which is generally executed in front or to the side of the robot. As the robot cannot raise its arms and simultaneously make a specific movement to the front, the emotional poses cannot be executed at the same time as the gestures.

6.4 Quiz

The activity which the child and the robot do together is a quiz. This quiz is an activity which the child and the robot play against each other. The robot will start by asking a child a question and providing four possible answers. These answers are projected on a tablet, which is then shown to the child. This way, the child can read the question and answers again before giving an answer. The robot will tell the child if the answer was correct, and ask the child to try again if this was not the case. The child can thus always try to give the correct answer two times. After this, the child gets the turn to ask a question. This question will be shown on the tablet screen and the child can read it to the robot. The robot will then give an answer. As we wish for the robot and the child to be equals in this game, the robot has a 0.25 chance of giving the wrong answer each time. After the robot has answered, the child can tell if this was the correct answer or not. The turn then goes back to the robot and the whole procedure starts again.

6.4.1 Questions

When a quiz is played, one of the key elements are of course the questions. The questions used in the experiment can be divided into trivia and general health questions. The trivia questions were developed in cooperation with children aged 9 to 11 and consist of a broad range of subjects such as food, animals, sports, music and television. The health questions were developed by (Zalm, 2011) and have been validated by a teacher and used in previous experiments. A list of all questions used can be found in Appendix 9.

¹<http://www.aldebaran-robotics.com>

Chapter 7

Experiment

7.1 Introduction

This chapter will report the experiment which has been done with the model of expressive behaviour as described in the previous chapters. In this experiment, children play the quiz with a robot that shows the model-based adaptive expressive behaviour and a robot without such a model. Four research questions are distinguished. First we would like to get an empirical foundation of the inner workings of this model. As this is a model designed to adapt a robots expressive behaviour to that of the child, we can ask the question of how the robots emotions follow that of the child. The question is: What is the influence of the child on the inner state of the robot? Based on this question, we can establish the following hypothesis.

1. The emotions of the robot will approach those of the child.

The emotions of the robot are not fully dependent on those of the child. The robot will also be influenced by occurrences emotionally relevant to itself. These occurrences can be both positive and negative and have their own arousal and valence values. Winning a game, for instance, is a positive occurrence with high arousal and valence values, while answering a question wrong is a negative occurrence with low arousal and valence values. The model should work in such a way that these relevant occurrences have a stronger influence on the robots emotions than the emotions of the child. The second hypothesis is the following.

2. Emotional occurrences will have a substantial effect on the emotions of the robot.

One final thing we would like to know about the model is in which way the robot compensates its emotions. When the valence or arousal of a child becomes too low, or the arousal too high, the

robot will not follow the emotions of the child but compensate. For a low valence, for instance, the robot will lower its valence when it is positive, but raise its valence when it is negative. In this experiment, we would like to know how this compensating works in practice.

In this experiment, the children will play a quiz with the robot. The quiz is designed in such a way that the robot will have a 75% chance each time to answer a question correctly. It is important, though, to check if this actually happens and to also consider the scores of the children. Although we do not expect the robot's expressive behaviour to influence the scores of the child, it would be good to know how difficult the quiz is for children.

Aside from knowing in which way the model and the quiz work, we also wish to find out what its influence is on the experience and (interaction) behaviour of the children. This question can be divided into two questions, one about the experience of the children and the other about their behaviour. When looking at experience, we wish to know if the expressive behaviour of the robot influences the opinion of children. When asked, do children have a preference for an expressive robot over a non-expressive robot? In case of the interaction behaviour, we wish to know if the children behave different with an expressive robot than with a non-expressive robot? More specifically, are they themselves more expressive or more positive with an expressive robot? Based on these questions, we can establish the third and fourth hypotheses.

3. When asked, children show a higher preference for an expressive robot than for a non-expressive robot.
4. Children behave more expressively and more positively with an expressive robot than with a non-expressive robot.

With these hypotheses in mind, an experiment was designed to test them. In the following sections, the methods of this experiment will be described in detail.

7.2 Experimental Method

In this section, the design and methods of the experiment will be described.

7.2.1 Experimental Design

We applied a within-subjects design with a two-level independent variable: the adaptive expressive behaviour of the robot, whether or not the robot showed the expressive behaviour. This means that two robots were used, one with and one without the expressive behaviour. With the third and fourth hypothesis, we can establish two dependent variables. These are the

preferences of the children and the behaviour of the children when interacting with the robot. Each of these variables can be further specified as described in the Measures section.

Inner workings of the model

Aside from the dependent variables described above, we would also wish to get a better understanding of the inner workings of the model. For this reason, the internal state of the robot at any point in time will be logged. With these data, we can get a better understanding of the relation between the emotional state of the child and that of the robot.

During the interaction, the robot will adopt the extroversion of the child as its own. In order to determine the extroversion of the child, the corresponding questions from the BFQ-C questionnaire were used. This questionnaire is validated for children (Muris, Meesters, & Diederens, 2005) and will give an insight to the extroversion of the children in the form of a score between 0 and 100.

7.2.2 Participants

All participants for this experiment were children from the primary school Dalton Lange Voren in Barneveld (group 5 and 6). 18 children participated in this experiment, mean age was 8.89, SD 0.81. 9 boys and 9 girls participated in the experiment. The mean extroversion of the children was 69, SD 10. This variance was taken as enough evidence that the BFQ-C questionnaire was not answered in a socially acceptable way. If all children would have had scores close together, this score could not have been used to adapt the robots behaviour to one child specifically. As this was not the case, however, the scores from this questionnaire could be used to establish the extroversion of the expressive robot.

7.2.3 Measures

The goal of this experiment was to test the effect of expressive behaviour of the robot on the experience and interaction (behaviour) of the child. Measuring experience and behaviour is, however, not straight-forward. For this reason, two kind of measures were used which give information about experience and behaviour. The first kind of measure used took the form of questionnaires filled in by the child. All except the emotion questions in these questionnaires have been successfully used in experiments on robot-child interaction before (Looije, Neerincx, & Lange, 2008; Robben, 2011). The full list of questions presented in all questionnaires can be found in Appendix 10. After playing with each robot, the children answered a number of questions about their experience with the robot. After playing with both robots, the children answered another questionnaire which was aimed at comparing the robots on the same points as the previous questionnaires. The answers from this questionnaire should give an insight as

Subject	Nr. of questions after each robot	Nr. of questions after both robots
Fun	9	1
Acceptance	3	1
Empathy	3	1
Trust	3	1
Emotions	3	1
Preference	0	1

TABLE 7.1: Topics of questions

to which robot they preferred and why. In Table 7.1, an overview of the different aspects of these questionnaires is given. Questionnaires with children are always subjective and prone to socially acceptable answers. Moreover, previous studies have shown that children tend to like every robot so much that not much difference between robots can be found with questionnaires. They simply all get a maximum score. For this reason, another kind of measure was used. All interactions between children and robot were recorded on video. With this video material, the expressive behaviour of the children during the interaction could be analysed. By counting specific expressions such as laughter and frowns, the expressiveness of the children was noted. This way it became possible to see if the children are more expressive or more positive when interacting with the expressive robot than when interacting with the non-expressive robot. The specific expressions noted and their definitions can be found in Table 7.2. To determine a score for the expressiveness of the child, the frequency scores of the behaviours were added up, counting smiles, frowns and startles once and laughter, bouncing, positive vocalization, shrugging, sighing and negative vocalization double. This was done to control for the fact that some expressions are stronger than others, e.g. laughter is a stronger expression than a smile. To determine the valence of the expressive behaviour, all frequencies of behaviours were used as scores and added up as shown in the formula below. This way, we get a balanced image of how positive or negative the child was. Note that frowns were left out of this score because they can denote either a positive (concentration), negative (misunderstanding) or neutral (thinking) child.

$$\text{Valence expressiveness} = \text{Smiles} + 2x(\text{Laughter} + \text{Bouncing} + \text{PosVocalization}) - (\text{Startle} + \text{NegVocalization}) - 2x(\text{Shrugging} \ \& \ \text{Sighing})$$

Expression	Properties	Value
Smiles	All instances where the mouth of the child angles upwards. As we only count instances and not duration, this was only counted when there was a change. So only when the mouth angles rose upwards.	+
Laughter	All cases in which the child laughed. Laughter is here classified as those smiles which are accompanied by sound or movement of the chest related to the happy feelings.	++
Excited bouncing	All cases in which the child either bounced up and down out of obvious excitement, or in which the child made a large excited gesture. An example of the latter is raising both arms, and other such gestures of success.	++
Positive vocalization	Every positive exclamation not directly related to the dialogue. Common words are <i>yay</i> or <i>yes</i> .	++
Frowns	All facial expressions obviously related to thinking, concentrating or misunderstanding. Also all facial expressions where the eyebrows are lowered.	0
Shrugging & Sighing	Raising the shoulders and dropping them again, or audibly letting out air. These two expressions are seen as signs of boredom	- -
Startle	All signs of involuntary fright from the child, such as it being startled by sudden movement.	-
Negative vocalization	All negative exclamations not directly related to the dialogue, such as <i>nou zeg</i> or <i>jammer</i> .	-

TABLE 7.2: Expressions and their definitions

7.2.4 Materials

The list of materials for this experiment can be divided into two categories, the technical devices and the computer programs. When it comes to the technical devices, two Nao robots were used, one Samsung Galaxy tab2 10.1 tablet on a seesaw developed for ALIZ-e, a Sony HDR-CX260VE camera, a Dell LATITUDE E6500 laptop and a TP-Link router. The laptop was used by the WoOz and ran the WoOz interface program through which the dialog was managed, the quiz operated and the emotional state of the child interpreted. It also ran the GOAL program which made the decisions on which behaviour to display in the way described in chapter 4 and 5. Because the two robots used are identical in appearance, both wore a different little shirt. One robot had a plain orange shirt, the other a striped white and orange shirt. These shirts were used to make sure that the children understood that there were two different robots and help them to keep the robots apart. In addition to keeping the robots apart, it was important that the children remember the names of the robots, as the questionnaires refer to them by the names *Charlie* and *Robin*. For this reason, both robots wore name badges on their shirts. Figure 7.1 shows the physical set-up of the experiment, Figure 7.2 shows the tablet as mounted on the seesaw and figure 7.3 shows the two robots with their shirts on.

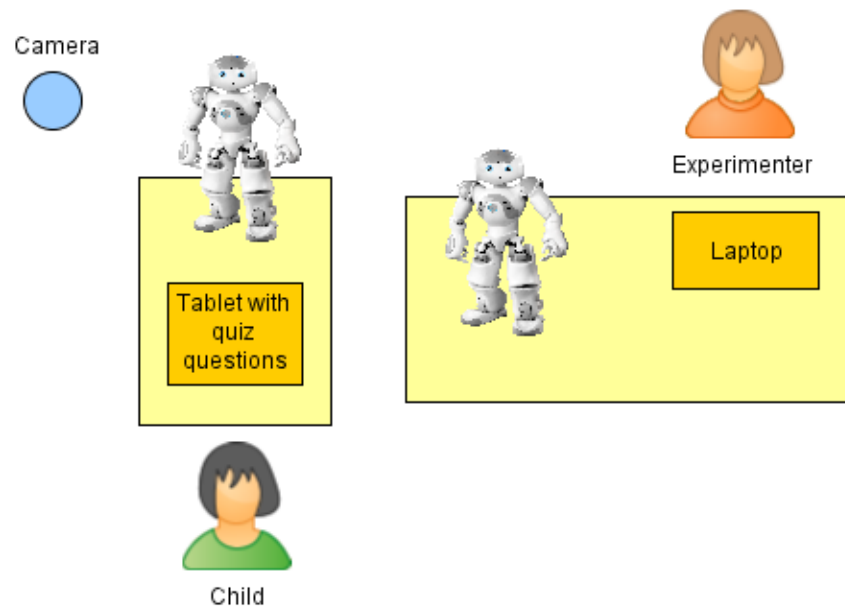


FIGURE 7.1: Set-up experiment



FIGURE 7.2: The robot turning the tablet on the seesaw to the child.

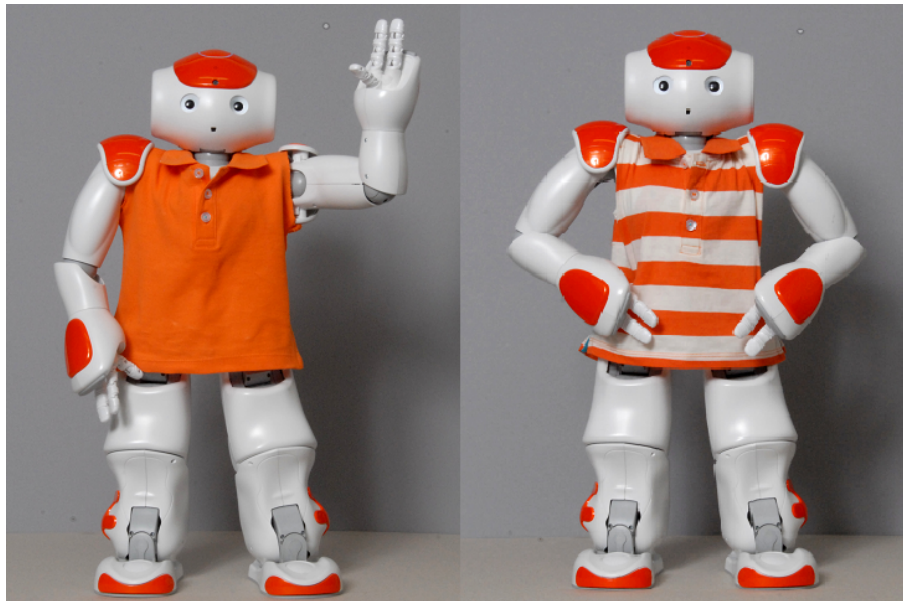


FIGURE 7.3: The robots used in the experiment with their shirts on.

7.2.5 Procedure

The experiment was conducted in two sessions, an introduction session and an experimental session. The introduction session was the same for all participants and took the form of a short classical lesson with the robots. In this lesson, one robot was introduced to the children in order to make them more familiar with robots. This robot did not wear a shirt and was given a different name than the robots used in the experimental sessions. After the introductions, all children filled in the BFQ-C questionnaire. In the experimental session, the first robot was always named Charlie and always used the same dialogue and questions, while the second robot was always named Robin and also always used the same dialogue and questions (different from the first robot, of course). Which robot used the expressive behaviour was counterbalanced, half of the children played the first quiz with the expressive robot, half with the non-expressive robot. The children were shown into the room and the experimenter first explained the quiz. In all sessions the first robot started with introducing itself to the child. After a short conversation about their interests, the robot asked if the child still understood the quiz and explained again when necessary. Next, the child and the robot played the quiz. After 12 questions (about 10 minutes), the robot ended the quiz and the interaction. The children were then presented with the questionnaire about the first robot. The first robot was then taken away, but kept in sight, and the second robot was brought to the child. The reason both robots are kept in sight is to ensure that the child views the robots as two different entities. Even though the robots wear different shirts, we do not wish for any suspicion regarding whether or not the robots are actually the same. The procedure described was repeated, the second robot introduced itself and had a short conversation with the child. The quiz was played for 10 minutes after which

Preparation session (whole class):
Children are introduced to a robot
Children take test about extroversion level
Experiment (each child individually):
Experimenter shows child where to sit and explains quiz
Robot 1 introduces itself
Robot 1 plays quiz with child
Child fills in questionnaire about Robot 1
Robots are changed
Robot 2 introduces itself
Robot 2 plays quiz with child
Child fills in questionnaire about Robot 2
Child fills in questionnaire about both robots
Picture with child and Robot

TABLE 7.3: Procedure experiment

the robot ended the interaction and the same questionnaire as before was presented. After this, one more questionnaire about the differences between the robots was presented. The session ended with the possibility for the child to take a picture with one of the robots. A schematic overview of the experimental sessions can be found in Table 7.3.

During the sessions, there are several points at which an emotionally relevant occurrence takes place. In the first dialogue, there is a happy occurrence when the quiz starts. During the quiz, every time a robot answers a question correctly there is a happy occurrence, every time the robot gives a wrong answer there is a sad occurrence. In the closing dialogue, there is a happy occurrence if the robot wins or ties with the child. Finally, there is a happy occurrence at the end if the child liked playing with the robot, a sad occurrence if the child did not. This adds up to either 6 or 7 emotional occurrences per interaction.

7.3 Results

The results of this experiment can be divided into three groups, the results on the inner workings of the model, the results on the preferences of the children and the results on the expressiveness of the children.

Inner workings of the model

From the log-data generated by the WoOz program, we can collect data about the inner state of the robot and the reported emotion child at each point in time. When we plot this data, we

	Child	Robot
Arousal Child 1	0.163	0.154
Valence Child 1	0.193	0.179
Arousal Child 2	0.108	0.184
Valence Child 2	0.143	0.197

TABLE 7.4: Standard Deviation in emotions of two children and the robot

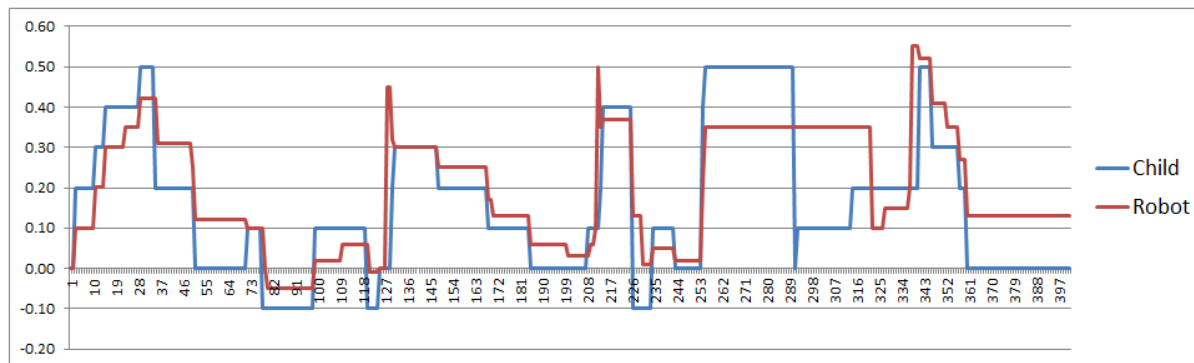


FIGURE 7.4: Arousal child 1 and robot

get a graph which shows the arousal or valence of the child and robot at each point in time. This graph visualizes in which way the emotions of the robot follow that of the child and shows when the robot was influenced by relevant emotional occurrences. Looking at the data from the different subjects we can distinguish two kinds of graphs. Figures 7.4 and 7.5 show a part of the valence and arousal graphs for a child with many changes in emotional stage, figures 7.6 and 7.7 a part of those of a child with few changes. Table 7.4 shows the standard deviations in the emotions of the children and the robot. The difference between the children is clear from these standard deviations, the child with many changes in emotional state has higher standard deviations than the child with few changes.

Looking at the graphs and standard deviations from these two children, it is very clear that the way in which the emotions of the robot follow that of the child depend on how much the emotions of the child fluctuate. If there are many changes in the emotional state of the child, the emotions of the robot will follow those of the child quite closely, their standard deviation being a little lower. When an emotional occurrence takes place, the robot will be influenced by this only for a short period of time. If there are few changes in the emotional state of the child, the robot will be less influenced by the child and the robot's standard deviation will be higher. From figure 7.6 and 7.7, it becomes clear that with a child with few changes in emotional state the robot will be influenced quite strongly by emotional occurrences. After a while, the emotions of the robot will move towards those of the child again.

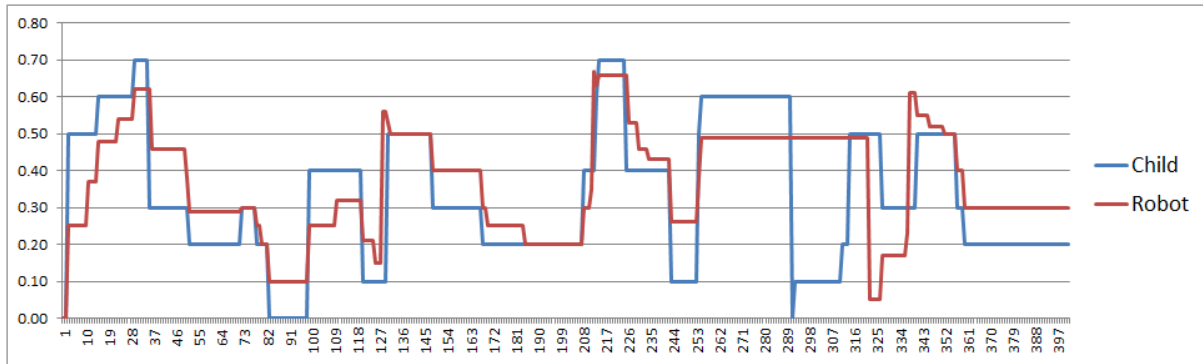


FIGURE 7.5: Valence child 1 and robot

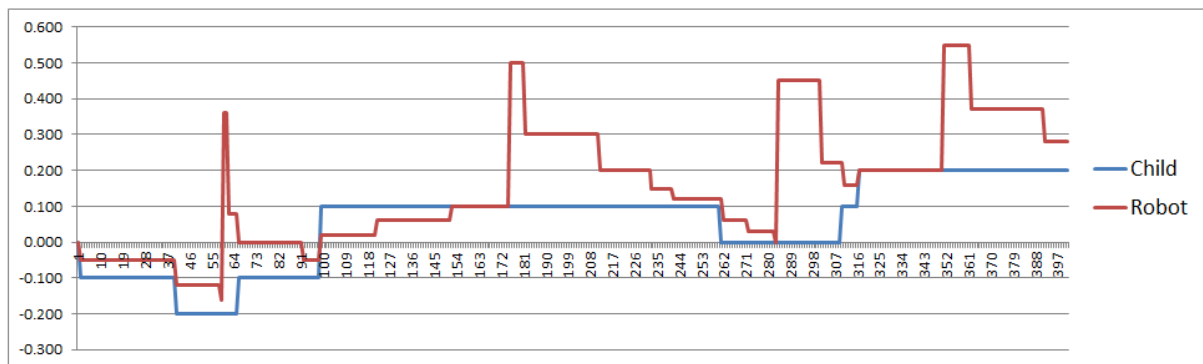


FIGURE 7.6: Arousal child 2 and robot

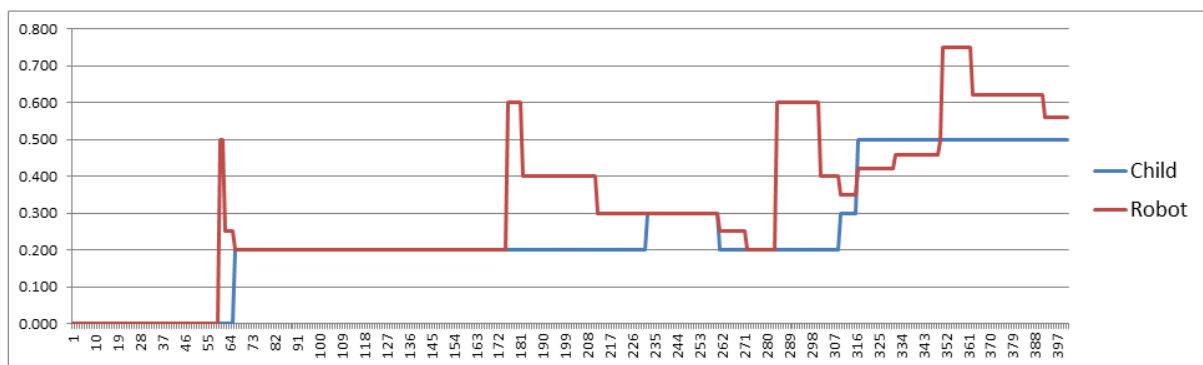


FIGURE 7.7: Valence child 2 and robot

	Child	Robot
Expressive Robot	90.74 %	69.44 %
Non-Expressive Robot	87.04 %	67.59 %
First Quiz	91.67 %	73.15 %
Second Quiz	86.11 %	63.89 %

TABLE 7.5: Results questionnaires about individual robots

The model also takes into account the situation in which the arousal and valence of the child become too low, or the arousal too high. Neither situation has occurred during this experiment, however. The low valence corresponds to a very negative emotion, the low arousal corresponds to sleepiness or boredom, and all children were very happy and excited to play with the robot. High arousal valued did occur, but the arousal will not be considered too high until it is above 0.9, which was never reached. We can therefore not conclude anything about the way in which the robot compensated its emotion.

Quiz scores

Table 7.5 presents the quiz scores in percentages. They are divided once into the scores with the expressive and the non-expressive robot per child, and once into the scores for the first and the second quiz per child. This last division is relevant as the quiz questions were different for the two quizzes played. For the expressive comparison, the children ($Mdn = 5$) scored significantly higher than the robot ($Mdn = 4$), $T = 1.50$, $p < 0.05$. In the non-expressive comparison, the children ($Mdn = 5$) also scored significantly higher than the robot ($Mdn = 4$), $T = 4.50$, $p < 0.05$. This also held for the first quiz, in which the children ($Mdn = 5.5$) scored significantly higher than the robot ($Mdn = 3.5$), $T = 3.50$, $p < 0.05$. Finally, also in the second quiz the children ($Mdn = 5$) scored significantly higher than the robot ($Mdn = 3.5$), $T = 2.50$, $p < 0.05$. There were no significant differences in the scores of the children for the expressive and the non-expressive robot, nor for the first and the second quiz.

Preferences of the children

Table 7.6 presents the scores from the questionnaires after each robot. The questions were asked on a scale from 1 to 5, meaning that a score of 100% corresponds to the most positive answer given to every question and a score of 0% to the most negative answer given to every question. Figure 7.8 further visualizes these scores and makes it clear that any differences found between the expressive and the non-expressive robot are minimal. None of the differences are statistically significant.

Subject	Expressive Robot	Non-Expressive Robot
Fun	89.12 %	87.96 %
Acceptance	85.19 %	86.57 %
Empathy	83.08 %	82.18 %
Trust	87.50 %	87.96 %
Emotions	75.00 %	75.00 %

TABLE 7.6: Results questionnaires about individual robots

Subject	Expressive Robot	Non-Expressive Robot
Fun	50.00 %	50.00 %
Acceptance	47.06 %	52.94 %
Empathy	55.88 %	44.12 %
Trust	38.24 %	61.76 %
Emotions	56.25 %	43.75 %
Choice	56.25 %	43.75 %

TABLE 7.7: Results final questionnaire comparing the robots

Table 7.7 and figure 7.9 show the results from the final questionnaire, comparing the two robots. Some data were excluded from this dataset, based on the motivations of the answers given. In one session, the program controlling the expressive robot crashed twice, meaning it had to be restarted. During this period of time, the robot did nothing. In the motivation section of the final questionnaire, the child preferred the non-expressive robot in every case, giving as motivation that it *was quicker to answer questions*. As this can be contributed to un-planned circumstances, the data was not taken into account. One other child motivated two of his choices for the non-expressive robot with reasons such as *because it showed more emotion*. When analysing the video material, it became obvious that this child had a difficult time keeping the names of the robots straight. These two factors indicate that he probably confused the robots in the final questionnaire. For this reason, his answers were not taken into account. Finally, one child preferred one robot over another because he liked its shirt and name more and stated this very clearly. This data was also excluded. When considering Table 7.7, note that the children had to choose between the robots, meaning that the scores for any subject will add up to 100%. All these scores are based on a single question. Although some differences can be seen, none are statistically significant. Finally, the children have motivated their answers to the final questionnaire. The top-3 reasons for choosing either robot can be found in Table 7.8.

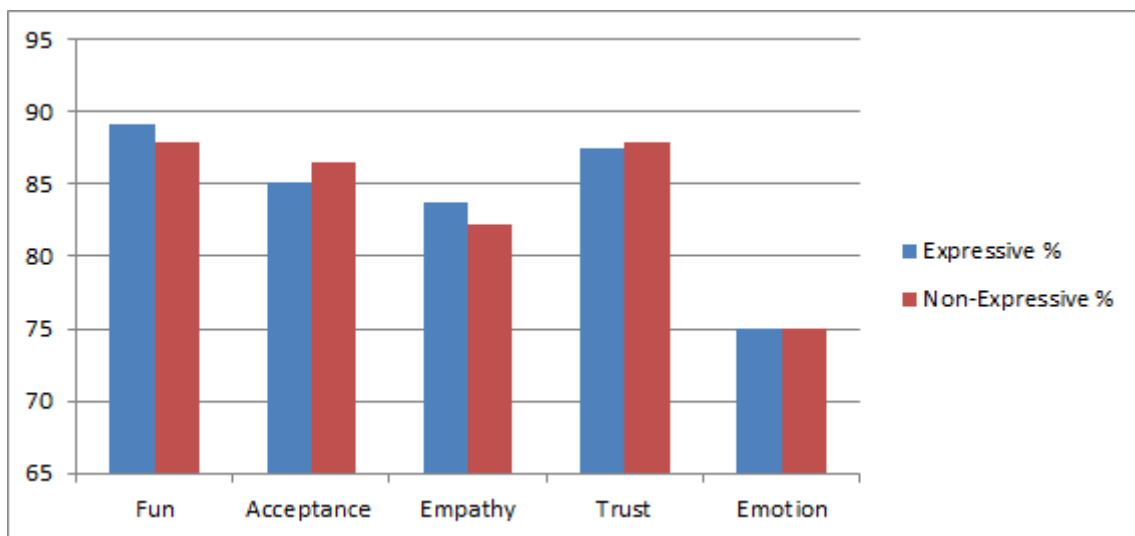


FIGURE 7.8: Results questionnaires about individual robots

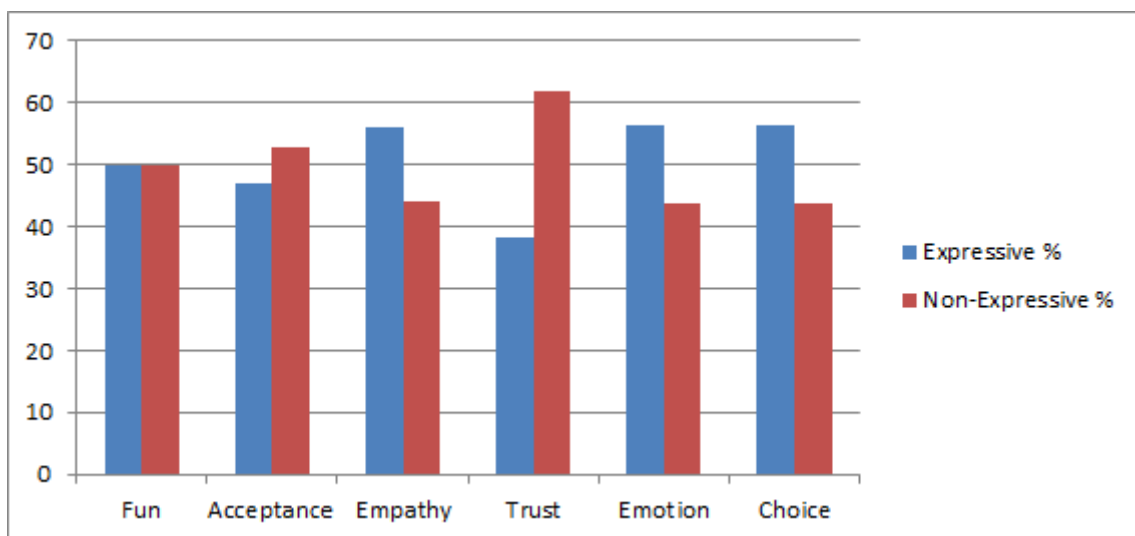


FIGURE 7.9: Results final questionnaire comparing the robots

Expressive Robot	Times given	Non-Expressive Robot	Times given
Emotion	13	Easy to understand	9
Movement	5	Reliable	6
Emotional poses	3	Calmer	2

TABLE 7.8: Reasons for choosing one robot over the other

Measure	Expressive Robot	Non-Expressive Robot
Expressiveness	33.59	29.24
Valence Expressiveness	29.06	24.94

TABLE 7.9: Results expressiveness analysis

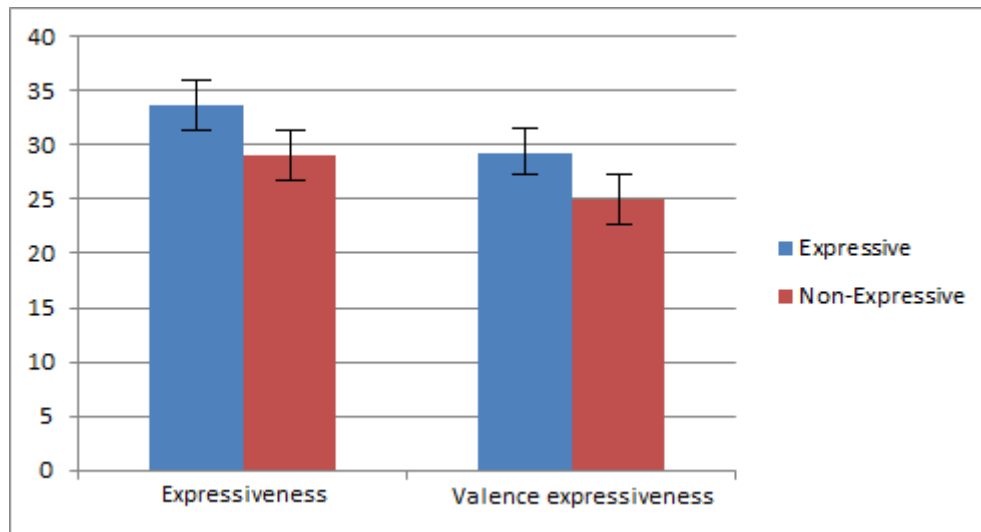


FIGURE 7.10: Results expressiveness analysis

Expressive behaviour of the children

The final set of results are those representing the expressiveness of the children during the interaction. In one session there was a technical problem with the camera, meaning that for one subject no video was available for analysis. The expressiveness was scored as described in section 2.3 (Measures) of this chapter, by the experimenter. In order to ensure the objectivity of this scoring method, two children were also scored by a second person with the instructions as stated in section 2.3. These results show that the differences between conditions are comparable. Table 7.9 shows the expressiveness scores of the children when interacting with the expressive and the non-expressive robot, Figure 7.10 visualizes these results. The results for overall expressiveness of the child for the expressive robot ($M=33.59$, $SE=17.34$) are significantly higher than for the non-expressive robot ($M=29.06$, $SE=13.53$), (t)(16)=2.156, (p) <0.05, (r)= 0.47 (one-tailed). Of course, we wish to make sure that the children did not express themselves more negatively with the expressive robot, so it is important to also consider how positively the children expressed themselves. These results show that the children had a significantly higher valence in their expressiveness with the expressive robot ($M=29.24$, $SD=16.75$) than with the non-expressive robot ($M=24.94$, $SD=13.89$) (t)(16)=2.251, (p) <0.05, (r)= 0.54(one-tailed).

Correlations

Aside from the results reported in the previous sections, a correlation analysis was done. The factors included in this analysis are the age of the child, its gender, their extroversion score, the differences between the expressive and non-expressive robot for each topic of the questionnaire, the robot of their choice in the final questionnaire, the number of times each robot was chosen in the final questionnaire, and the differences between both expressiveness and valence of the expressiveness of the children. The full correlation matrix can be found in Appendix 11.

7.4 Discussion and Conclusions

With the results from the previous section, we can provide some insight in the hypotheses posed at the beginning of this chapter.

Inner workings of the model

The two hypotheses about the inner workings of the model were:

1. The emotions of the robot will approach those of the child.
2. Emotional occurrences will have a substantial effect on the emotions of the robot.

Looking at the graphs in figures 7.4, 7.5, 7.6 and 7.7, we can see that the extent to which these hypotheses were confirmed depends on the emotional behaviour of the child. Figures 7.4 and 7.5 show the emotions of a child which has many fluctuations in emotional state. Both the arousal and valence of this child changed quite often. For this child, the first hypothesis can be confirmed. The arousal and valence of the robot clearly follow those of the child. The confirmation of the second hypothesis is not quite as clear, however. The emotions of the robot sometimes peak over or drop below those of the child (indicating a happy or sad occurrence), but these moments can only be spotted when looking closely. The influence of emotional occurrences do not, therefore, have a visible strong effect of the emotions of the robot. The second hypothesis cannot be confirmed when looking at the data from this child. When considering figures 7.6 and 7.7, however, we get a different image. These graphs show the arousal and valence of a child with very few fluctuations in emotion. In these graphs, we can clearly see four peaks in the emotions of the robot, which correspond to happy occurrences. For this child, the second hypothesis can therefore be confirmed. Considering the first hypothesis for this child, we can see that the emotions of the robot slowly approach those of the child after each emotional occurrence. We can therefore also confirm the first hypothesis for this child. Concluding, we can state that the emotions of the robot will approach those of the child with this model. Emotional occurrences will only have a visible strong effect when there are few changes in the emotional state of the child.

Quiz Scores

From the results in the quiz we can see that the children are significantly better at the quiz than the robot. Although it is not a bad thing if the child is slightly better at the quiz than the robot, we would not want the quiz to become too easy. These results mean that the questions for the quiz could be made a little more difficult for children from 8 to 10 years old. We can also conclude that the expressive behaviour of the robot did not have any influence on the quiz-scores of the children.

Opinions of the children

The third hypothesis was about the effect of the expressive behaviour of the robot on the opinions of the children.

3. When asked, children show a higher preference for an expressive robot than for a non-expressive robot.

Looking at the results, we see that when asking the children for their opinions of each robot, no significant differences can be found between the expressive and non-expressive robot. We cannot, therefore, confirm the hypothesis and state that children show a higher preference for an expressive robot than for a non-expressive robot. It is possible, however, to make some suggestions when combining the data from the final questionnaire with the motivations given to the answers. Interesting from figure 7.9 is that although the expressive robot scores higher on empathy, emotion and general preference, the non-expressive scores higher in acceptance and trust. These differences may be due to chance as none are statistically significant, but when looking at the motivations to these questions we can see some trends. Table 7.8 gives the top 3 kind of answers for choosing either the expressive or the non-expressive robot, for any of the questions. Looking at these reasons, we see that children particularly like the fact that the expressive robot showed its emotions and that it moved more. These reasons are given most often for the questions about fun, empathy and emotion. For the non-expressive robot, the strongest argument for choosing it was that it was easier to understand and more reliable. These reasons were given most to the questions about fun, acceptance and trust. If we take these motivations as predictors we could state the following hypotheses. We would expect the expressive robot to score highest on the empathy and emotion scores and the non-expressive robot to score highest on the trust and acceptance scores. Going back to the data, we can see that this corresponds to what we would expect from the motivations. This means that although we cannot conclude that children have a preference for an expressive robot, there is some reason to believe that an expressive robot increases empathy and anthropomorphism, but decreases acceptance and trust. The image which arises from the motivations to answer is that for some children, the showing of emotions was very important in determining their preference for the expressive robot. When asked which robot they thought nicer, one girl motivated her choice for

the expressive robot with *She showed her feelings and because of this I felt a stronger friendship*. This motivation gives a very clear statement of the positive effect showing emotion can have on robot-child interaction. There is also a downside to the expressive behaviour, however. Especially the voice of the expressive robot has proven to make the robot harder to understand. The questionnaires show that it is very important for children to have a robot which they can understand well. One child which preferred the expressive robot in every other question, choose the non-expressive as being easier to use because it *speaks more clearly*. We can conclude that intelligibility is very important to children than emotion when it comes to a robots voice.

Behaviour of the children

The final hypothesis was about the expressive behaviour of the children.

4. Children behave more expressively and more positively with an expressive robot than with a non-expressive robot.

When looking at expressiveness scores for children in Table 7.9 and Figure 7.10 we can clearly see that children behave more expressively with an expressive robot than with a non-expressive robot. Moreover, when controlling for expressions being either positive or negative, we also see that children behave more positively in their expressions with an expressive robot than with a non-expressive robot. We can therefore confirm this hypothesis. Although there are very large differences in expressiveness scores between children, the expressive robot tends to incite more smiles, more laughs, etc. from children. Even though the relation between how much one positively expresses oneself with a robot and how much one likes that robot is not fully clear, positive expressions are a sign of enjoyment. We can, therefore, state that children enjoyed themselves more with an expressive than with a non-expressive robot.

Chapter 8

Conclusion

8.1 Conclusion

Combining the results from the previous chapters it is possible to give some answers to the research questions. The two research questions posed at the beginning of this thesis are:

- How can we make the Nao use expressive behaviour in robot-child interaction?
- What is the effect of the Nao using the model for expressive behaviour in interaction with children?

We can give two kinds of answers to the first research question, a theoretical and a practical answer. Based on a review of literature from the fields of psychology and social robotics, we can draw several conclusions on what kind of behaviour the Nao robot should show and when.

The expression of emotions has an important role in interactions, for instance suppressing one's emotion has a negative effect on the forming of bonds between people. The Nao robot should, therefore, show its own emotions. Firstly, these are influenced by occurrences, e.g. winning a game will make the robot happy, losing will make it sad. People both mimic and influence the emotions of others when they interact, so the Nao robot should also be influenced by the emotions of the child. Emotion can be conveyed and recognized well through both body posture and voice. Full body poses of anger, sadness and happiness are particularly well recognized in humans and these poses have also successfully been developed for the Nao. As the Nao cannot execute these non-stop, however, it should only use these poses when something particularly relevant occurs. The position of the trunk of the body seems closely related to valence in particular, while other research shows that head position of a robot influences both perceived arousal and valence. The robot should thus adapt its trunk and head position based on its

valence and arousal. Research into emotional speech shows that high-arousal emotions have a higher fundamental frequency, energy and speech rate. The robot should therefore adapt its voice based on arousal. Spontaneous gestures are the type of gestures most commonly used in speech and can be divided into different types. Which type of gesture occurs depends on the type of narrative it corresponds to. The robot should thus choose its gestures based on its speech and the type of narrative.

Although these conclusions give us an image of how the robot uses expressive behaviour, we also wish to know how we can make the robot do this. Chapter 5 shows how we can use the BDI language GOAL to implement the expressive behaviour of the robot. The input for this implementation are the emotions of the child, the arousal and valence of emotional occurrences and the possible gestures to choose from. When the model receives a list with gestures to choose from, it will make a choice based on the type of text the gesture will correspond with and the type of gesture. The specific gesture chosen is output and is sent to the robot. When information about emotions enter the program, the implementation will first adapt the internal parameters of arousal and valence of the robot based on the input. From these parameters, the program will know how to adapt the behavioural values of the robot, such as head position, speech rate, etc. These adaptations in behaviour are the output of the implementation and will be sent to the Nao robot.

What is the effect of Nao using the architecture for expressive behaviour in interaction with children?

In order to answer this question, an experiment with the Nao robot displaying expressive behaviour was done. The first effect measured was the way in which the emotions of the robot related to those of the child. From these results we can conclude that the emotions of the robot follow those of the children. If the child has a lower arousal than the robot, the arousal of the robot will go down and the other way around. The only times where the emotions of the robot do not approach those of the child is when an emotionally relevant occurrence takes place. These emotional occurrences only have a visible effect on the emotional state of the robot, however, if the emotional state of the child does not fluctuate too much.

The second effect measured in this experiment was the subjective opinion of the children. These opinions were measured with questionnaires. From the results we can conclude that there is no statistically significant difference in the opinions of children about an expressive and a non-expressive robot. A part of the questionnaires also asked for motivation for the choices of the children. Based on these motivations, we can draw several conclusions about the opinion of the children. The most negative point of the expressive behaviour of the robot was its speech. The way emotions were displayed through the voice made the robot harder to understand. This undermined the acceptability of the robot. The children also trusted the non-expressive robot

more than the expressive robot. From these results we can conclude that the adaptations to the speech of the robot need to be carefully revised, as they undermine the children liking the expressive robot. The most positive point of the expressive behaviour of the robot were its emotions and its movement. These factors made the children feel more empathy towards the robot. One child even stated that she thought the expressive robot to be nicer because *she showed her emotions and because of this I felt more friendship*. From this statement we can conclude that showing emotion is a very promising way of improving empathy and friendship between robots and children.

The final effect measured was the expressive behaviour of the children. An analysis of the recorded interaction shows that the children behave significantly more expressive with an expressive robot than with a non-expressive robot. In human interaction expressive behaviour, in particular showing emotion, elicits expressive behaviour from the interaction partner. This means that at the greater expressiveness of the children with an expressive robot at the very least means that the children respond to expressive behaviour. Moreover, this could suggest that the child interpreted the behaviour of the robot in the same way it would interpret a human, and that the behaviour of the robot was, therefore, natural. This is, however, quite a claim. Stronger conclusions can be drawn looking at how positively the children expressed themselves. The results from the experiment show that the children also expressed themselves more positively with the expressive robot than with the non-expressive robot. I will adopt the assumption that when children express something positive (such as a smile), they are enjoying themselves. We can, therefore, conclude that children enjoy themselves more when interacting with an expressive robot than with a non-expressive robot.

Conclusion

Summarizing the conclusions from the two research questions, we can conclude that the Nao robot should display expressive behaviour in voice and body based on its emotions and its speech. Its emotions should be influenced by the child it interacts with, as well as by relevant occurrences such as winning a game. An experiment has shown that children particularly like if a robot shows emotion through movement, while showing emotion through voice has the negative effect of reducing intelligibility. Moreover, children enjoy interacting with an expressive robot more than interacting with a non-expressive robot.

8.2 Implications and Further Research

The conclusions drawn have some implications for future work in robot-child interaction. First, we see from the reactions of the children to the robot that the emotional poses designed by

Beck, Hiolle, et al. (2010) and Aldebaran are well recognized by the children. Especially the happy poses elicited positive reactions from the children. This indicates that these poses can be well used within a broader model of emotion expression. A second implication is related to the questionnaires used in the experiment. When previously used by Robben (2011), the differences in scores between children turned out to be quite small, which led to the theory that the children might have given socially acceptable answers. In this study, the differences were actually much larger, which indicates that the low variation in scores was probably due to chance. This is a good thing, as it increases the usability of the BFQ-C questionnaire in adapting a robot's personality to that of a child. The results from the other questionnaires, or rather the lack of results, is a further confirmation that questionnaires are not the best way to compare opinions of children on robots. The motivations given to the questions in the final questionnaire have, however, proven to be very useful. Moreover, the questionnaire comparing two robots got stronger results than the questionnaire per robot. From this, we can draw the conclusion that questionnaires comparing two robots should be preferred over questionnaires about single robots and that asking for motivations to questions can give essential insight into the results. Finally, a very important implication of the conclusion is that it can be very rewarding to pay attention to the behaviour of children as well as to their opinions. An analysis of behaviour may result in important insights which are not gained through subjective measures.

Although some progress has been made in developing expressive behaviour, much work also still remains to be done. Firstly, several suggestions can be made to further improve or understand the model for expressive behaviour. The first thing which is still unclear is the role of emotion expression through voice in robot-child interaction. Although people successfully display emotion in their voice, the result of the robot doing so was unintelligibility. This corresponds to the findings of Kessens et al. (2009), who found that although children understand emotional speech, they find it harder to understand. What exactly needs to be improved is not clear yet. It might be that the robot spoke too fast or that the pitch was too high. A recent study by Van Dam (2013) with an extrovert robot seems to contradict this, though. The extrovert Nao robot in this study spoke with a high pitch and speech rate, using the same text to speech generator as was used in this thesis. None of the children liked the robot any less because of its voice. This suggests that it might be the fluctuations in voice which made the robot harder to understand, because the voice was unpredictable. Finally, the problem might also be solved through improving the overall voice of the robot. One hypothesis is that when the overall intelligibility of the robot voice is better, showing emotion will not have such a negative effect anymore because the intelligibility will stay above a certain threshold. Another possible study on the voice of the robot would be if children recognize the emotions from the voice of the robot alone. Although comparable research has been done with success (Kessens et al., 2009), no study so far has used arousal to directly influence fundamental frequency, volume and speech rate. It would therefore be interesting to see if children recognize emotion from voice

alone. A second factor of the model which is not fully clear yet is the expression of emotion through general posture and eye colour. Although the children were very positive about the robot showing its emotions, most didn't specify if they recognized the emotion from specific body poses, or also from general posture. It would be interesting to find out if head position, trunk position and eye colour alone can be used to convey emotion to children. Eye colour has been used before with the Nao robot by Cohen et al. (2011) to support emotion expression, but in this study it was coupled with full body poses. We therefore do not know if eye colour alone could convey emotion to children, or if it improves emotion recognition. Head position in the Nao robot has been studied before by Beck, Canamero, and Bard (2010), but not with children. We would, however, expect similar results for children as research has suggested that by the age of 8, children recognize emotions from body movement as well as adults (Boone & Cunningham, 1998).

Aside from the model itself, there are still some gaps when it comes to developing a robot capable of fully autonomous expressive behaviour. The experiment described in this thesis used a WoOz program to assist the robot. Most important for the model was that the experimenter interpreted the emotions of the child for the robot. This is, of course, something which should in the future be handled by the robot itself. Although this topic has not been covered in this thesis due to time constraints, it would be very interesting to see how the model would perform when a fully objective entity (the robot) interprets the emotions of the children. Another topic not covered is the automatic generation of gestures. The model relies on input of a set of ready-made gestures for each part of speech. Although it is possible at this point to hard-code all gestures, the development of automatically generated speech will change this. This means that in the future, gestures will have to be generated automatically from text. Several systems capable of doing this have already been developed (Cassell et al., 2001; Giorogolo, 2010; Aly & Tapus, 2011, 2012). The methods and results from these studies may prove very useful in generating gestures from text.

Finally, it would be interesting to further develop the method of scoring expressivity of children in order to evaluate robot behaviour. At this point, the exact relation between expressivity, valence of expressivity and opinions of children is not quite clear. One process which might play a role in this relation is emotion regulation, in which we influence how we express our emotions (J. J. Gross et al., 2011). Another aspect which should be taken into account is the fact that the way in which we express emotion is influenced by the familiarity of the interaction partner and the kind of emotion experienced (Kleck et al., 1976; Fridlund, 1991; Yarczower & Daruns, 1982; Buck et al., 1992). Further work might be able to clear up some of these relations, which would mean that stronger conclusions can be drawn based on this measure.

Glossary

back-channelling Utterances spoken by listeners during a narrative to show that they're attentive to what the other person is saying.

beat gestures Spontaneous gestures relating to the rhythm of speech, looking like beating to music.

biologically inspired robot A robot where the designers aim to create a robot which simulates social behaviour internally.

body language Expression through bodily posture and movement.

deictic gestures Spontaneous pointing gestures.

Duchenne laughter Laughter with as cause a genuine positive emotion.

emblematic gestures Culturally specific gestures which are consciously made, such as the thumbs up sign.

emotion A subjective state of feeling which involve arousal (either high or low), have a specific valence (ranging from very positive to very negative) and have characteristic forms of expression.

emotion regulation The process in which we influence which emotions we experience, when we experience them and how we express them. (J. J. Gross et al., 2011).

extranarrative clause Those clauses during narration which are not plot-related, such as descriptions.

facial expression The expression of emotion by means of facial features such as raising the eyebrows..

functionally designed robot A robot where the design objective is to create a robot which outwardly appears to be socially intelligent.

fundamental frequency The lowest frequency of a periodic waveform.

iconic gestures Spontaneous gestures which have a direct relationship to a concrete event as described or mentioned in speech.

metaphorical gestures Spontaneous gestures which have a direct metaphorical relationship to an abstract concept as described or mentioned in speech.

mimicry Copying the emotional expression of the people we're interacting with.

narrative clause The clause during narration which describe the plot.

preparation phase The first phase in making a gesture, the hand(s) move from their resting place.

propositional gestures Consciously made gestures with a direct relationship to what is being said.

retraction phase The final phase of making a gesture, the hand(s) goes back to the resting place.

rheme Those parts of an utterance which are new or interesting.

sociable robot A robot which has its own internal states and goals and interacts in a human-like way not only to the benefit of the interaction partner, but also to its own.

social interface robot A robot which uses human-like social cues and communication modalities.

social robot Those robots to which people apply a social model in order to understand and interact with them (Breazeal, 2003b).

socially embedded robot A robot which is situated in a social environment where it interacts with people and robots, which is structurally coupled with its social environment and is at least partially aware of interactional structures as used by people (Dautenhahn et al., 2002).

socially evocative robot A robot designed to encourage people to attribute human characteristics to it, but whose design goes no further than this.

socially intelligent robot A robot which shows aspects of a human-like style of social intelligence, which is based on models of human cognition and social competence (Dautenhahn, 1998).

socially receptive robot A robot which don't just interact in a human-like way, but also benefits from interactions with people itself.

socially situated robot A robot which is surrounded by an environment which it perceives and reacts to (Dautenhahn et al., 2002).

stance Refers to the aspect of emotion which determines if you want to approach the thing causing the emotion or move away. Fear, for instance, has a very low stance.

stroke phase The most important phase in making a gesture, this is the phase where the specific gesture takes place.

References

- Aly, A., & Tapus, A. (2011). Speech to head gesture mapping in multimodal human-robot interaction. In *Proceedings of the European Conference on Mobile Robotics*.
- Aly, A., & Tapus, A. (2012). Prosody-driven robot arm gestures generation in human-robot interaction. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction* (pp. 257–258). New York, NY, USA: ACM.
- Anderson, C., Keltner, D., & John, O. (2003). Emotional convergence between people over time. *Journal of Personality and Social Psychology*, *84*(5), 1054-1068.
- Arkin, R., Fujita, M., Takagi, T., & Hasekawa, R. (2003). An ethological and emotional basis for human-robot interaction. *Robotics and Autonomous Systems*, *42*, 191-201.
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614-636.
- Bavelas, J. B., Black, A., Charles, R., & Mullett, J. (1986). I show how you feel: Motor mimicry as a communicative act. *Journal of Personality and Social Psychology*, *50*(2), 322-329.
- Beck, A., Canamero, L., & Bard, K. (2010). Towards an affect space for robots to display emotional body language. In *IREE RoMan Conference*.
- Beck, A., Canamero, L., Damiano, L., Sommavilla, G., Tesser, F., & Cosi, P. (2011). Children interpretation of emotional body language displayed by a robot. In *ICSR 2011*.
- Beck, A., Hiolle, A., Mazel, A., & Canamero, L. (2010). Emotional body language displayed by robots. In *Affine 2010*.
- Beck, A., Stevens, B., Bard, K., & Canamero, L. (2012). Emotional body language displayed by artificial agents. *ACM transactions on Interactive Intelligent Systems. Special issue on Affective Interaction in Natural Environments*, *2*, 1-29.
- Benítez Sandoval, E., & Penalzoa, C. (2012). Children’s knowledge and expectations about robots: a survey for future user-centered design of social robots. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction* (pp. 107–108). New York, NY, USA: ACM.
- Beran, T., Ramirez-Serrano, A., Kuzyk, R., Fior, M., & Nugent, S. (2011). Understanding how children understand robots: Perceived animism in child-robot interaction. *International Journal of Human-Computer Studies*, *69*, 539-550.

- Boone, T., & Cunningham, J. (1998). Children's decoding of emotion in expressive body movement: The development of cue attunement. *Developmental Psychology, 34*(5), 1007-1017.
- Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62*(4), 645-657.
- Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology, 77*(3), 343 - 352.
- Boyatzis, C. J., & Varghese, R. (1994). Children's emotional associations with colors. *Journal of Genetic Psychology, 155*(1), 77-85.
- Breazeal, C. (2001). *Designing sociable robots* (C. Breazeal, Ed.). MIT P.
- Breazeal, C. (2002). *Designing sociable robots*. MIT Press.
- Breazeal, C. (2003a). Social interactions in HRI: The robot view. *IEEE Transactions on Systems, Man, and Cybernetics, 34*(2), 181-186.
- Breazeal, C. (2003b). Toward sociable robots. *Robotics and Autonomous Systems, 42*, 167-175.
- Brody, L. R. (1985). Gender differences in emotional development: A review of theories and research. *Journal of Personality, 53*(2), 102-149.
- Buck, R., Loscow, J., Murphy, M. M., & Costanzo, P. (1992). Social facilitation and inhibition of emotional expression and communication. *Journal of Personality and Social Psychology, 63*(3), 962-968.
- Burkhardt, F., & Stegmann, J. (2009). Emotional speech synthesis : Applications, history and possible future. In *Proc. ESSV*.
- Butler, E. A., Egloff, B., Wilhelm, F. H., Smith, N. C., Erickson, E. A., & Gross, J. J. (2003). The social consequences of expressive suppression. *Emotion, 3*(1), 48-67.
- Canamero, L., & Fredslund, J. (2001). I show you how I like you - can you read it in my face? *IEEE Transactions on Systems MAN and Cybernetics, 31*(5), 454-459.
- Cassell, J. (2000). Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. In J. Cassell, J. Sullivan, S. Prevost, & E. Churchill (Eds.), *Embodied conversational agents* (p. 29-63). MIT Press.
- Cassell, J., Vilhjlmsson, H. H., & Bickmore, T. (2001). Beat: the behavior expression animation toolkit. In *SIGGRAPH 2001*.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect : The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893-910.
- Chidambaram, V., Chiang, Y.-H., & Mutlu, B. (2012). Designing persuasive robots : how robots might persuade people using vocal and nonverbal cues. In *Proceedings of the seventh annual acm/ieee international conference on human-robot interaction* (pp. 293-300). New York, NY, USA: ACM.
- Cohen, I., Looije, R., & Neerinx, M. (2011). Child's recognition of emotions in robots' face and body. In *HRI 2011 Conference*.

- Colletta, J.-M., Pellenq, C., & Guidetti, M. (2010). Age-related changes in co-speech gesture and narrative: Evidence from french children and adults. *Speech Communication*, *52*(6), 565-576.
- Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions and viewpoint dependence. *Journal of Nonverbal Behavior*, *28*, 117-139.
- Dael, N., Mortillaro, M., & Scherer, K. (2012). The Body Action and Posture coding system (BAP): Development and reliability. *Journal of Nonverbal Behavior*, *36*, 97-121.
- Dautenhahn, K. (1998). The art of designing socially intelligent agents - science, fiction and the human in the loop. *Applied Artificial Intelligence Journal, Special Issue on Socially Intelligent Agents*, *12*, 12-7.
- Dautenhahn, K., Ogden, B., & Quick, T. (2002). From embodied to socially embedded agents : Implications for interaction-aware robots. *Cognitive Systems Research*, *3*(3), 397 - 428.
- De Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior*, *13*, 247-268.
- Digman, J. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*, 417-440.
- Ekman, P. (1965). Differential communication of affect by head and body cues. *Journal of Personality & Social Psychology*, *2*(5), 726-735.
- Ekman, P. (1972). Universals and cultural differences in facial expressions of emotions. In *Nebraska symposium on motivation* (p. 207-283). University of Nebraska Press.
- Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, *42*, 143-166.
- Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology*, *60*(2), 229-240.
- Giorogolo, G. (2010). *Space and time in our hands*. Unpublished doctoral dissertation, Universiteit Utrecht.
- Gross, J. J., Sheppes, G., & Urry, H. L. (2011). Cognition and emotion lecture at the 2010 SPSP emotion preconference, emotion generation and emotion regulation: A distinction we should make (carefully). *Cognition and Emotion*, *25*(5), 765-781.
- Gross, M., Crane, E., & Fredrickson, B. (2010). Methodology for assessing bodily expression of emotion. *Journal of Nonverbal Behavior*, *34*, 223-248.
- Guidetti, M. (2002). The emergence of pragmatics: forms and functions of conventional gestures in young french children. *First Language*, *22*(3), 265-285.
- Häring, M., Bee, N., & André, E. (2011). Creation and evaluation of emotion expression with body movement, sound and eye color for humanoid robots. In *RO-MAN, 2011 IEEE*.
- Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, *40*(2), 129-141.

- Hirth, J., Schmitz, N., & Berns, K. (2011). Towards social robots: Designing an emotion-based architecture. *International Journal of Social Robotics*, 3, 273-290.
- Hirth, J., Schmitz, N., & Berns, K. (2012). Playing tangram with a humanoid robot. In *Proceedings of ROBOTIK*.
- Hortacsu, N., & Ekinci, B. (1992). Children's reliance on situational and vocal expression of emotions: Consistent and conflicting cues. *Journal of Nonverbal Behavior*, 16(4), 231-247.
- Jung, S., Lim, H. taek, Kwak, S., & Biocca, F. (2012). Personality and facial expressions in human-robot interaction. In *Proceedings of the seventh annual acm/ieee international conference on human-robot interaction*.
- Kanda, T., Sato, R., Saiwaki, N., & Ishiguro, H. (2007). A two-month field trial in an elementary school for long-term human-robot interaction. *IEEE Transactions on Robotics*, 23, 962-971.
- Kaya, N., & Epps, H. H. (2004). Relationship between color and emotion: a study of college students. *College Student Journal*, 38(3), 396-405.
- Kessens, J. M., Neerincx, M. A., Looije, R., Kroes, M., & Bloothoof, G. (2009). Facial and vocal emotion expression of a personal computer assistant to engage, educate and motivate children. In *IEEE 2009*.
- Kim, A., Kum, H., Roh, O., You, S., & Lee, S. (2012). Robot gesture and user acceptance of information in human-robot interaction. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction* (pp. 279-280). New York, NY, USA: ACM.
- Kleck, R. E., Vaughan, R. C., Cartwright-Smith, J., Vaughan, K. B., Colby, C. Z., & Lanzetta, J. (1976). Effects of being observed on expressive, subjective, and physiological responses to painful stimuli. *Journal of Personality and Social Psychology*, 34(6), 1211-1218.
- Kolb, B., Wilson, B., & Taylor, L. (1992). Developmental changes in the recognition and comprehension of facial expression: Implications for frontal lobe function. *Brain and Cognition*, 20(1), 74-84.
- Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: Expression, experience and physiology. *Journal of Personality and Social Psychology*, 74(3), 686-703.
- Leite, I., Castellano, G., Pereira, A., Martinho, C., & Paiva, A. (2012). Modelling empathic behaviour in a robotic game companion for children: an ethnographic study in real-world settings. In *Proceedings of the seventh annual ACM/IEEE International conference on Human-Robot Interaction* (pp. 367-374). New York, NY, USA: ACM.
- Likowski, K. U., Mahlberger, A., Seibt, B., Pauli, P., & Weyers, P. (2008). Modulation of facial mimicry by attitudes. *Journal of Experimental Social Psychology*, 44(4), 1065 - 1072.
- Looije, R., Neerincx, M., & Hindriks, K. (2012). *How to develop a theoretical and empirical founded set of user requirements and applications for social robots*. (To Appear)

- Looije, R., Neerinx, M., & Lange, V. de. (2008). Childrens responses and opinion on three bots that motivate, educate and play. *Journal of Physical Agents*, 2(2).
- McHugo, G. J., Lanzetta, J. T., & Bush, L. K. (1991). The effect of attitudes on emotional reactions to expressive displays of political leaders. *Journal of Nonverbal Behavior*, 15(1), 19-41.
- McHugo, G. J., Lanzetta, J. T., Sullivan, D. G., Masters, R. D., & Englis, B. G. (1985). Emotional reactions to a political leader's expressive displays. *Journal of Personality and Social Psychology*, 49(6), 1513-1529.
- McNeill, D. (1992). *Hand and mind*. The University of Chicago Press.
- Muris, P., Meesters, C., & Diederer, R. (2005). Psychometric poperties of the big five questionnaire for children (bfq-c) in a dutch sample of young adolescents. *Personality and individual differences*, 38(8), 1757 - 1769.
- Nourbakhsh, I. (1999). An affective mobile robot educator with a full-time job. *Artificial Intelligence*, 144(1-2), 95-124.
- Nowlis, V., & Nowlis, H. H. (1956). The description and analysis of mood. *Annals of the New York Academy of Sciences*, 65(4), 345-355.
- Peters, K., & Kashima, Y. (2007). From social talk to social action: Shaping the social triad with emotion sharing. *Journal of Personality and Social Psychology*, 93(5), 780-797.
- Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, 89(4), pp. 344-350.
- Poldrack, R. A., Wagner, A. D., Ochsner, K. N., & Gross, J. J. (2008). Cognitive emotion regulation insights from social cognitive and affective neuroscience. *Current Directions in Psychological Science*, 17(2), 153-158.
- Reeves, B., & Nass, C. (1996). *The media equation*. CSLI Publications.
- Robben, S. (2011). *It's nao or never! facilitate bonding between a child and a social robot: Exploring the possibility of a robot adaptive to personality*. Unpublished master's thesis, Radboud Universiteit Nijmegen.
- Russell, J. (1980). A circumplex model of affect. *Journal of Personality & Social Psychology*, 39(6), 1161-1178.
- Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11(3), 273 - 294.
- Saerbeck, M., Schut, T., Bartneck, C., & Janse, M. D. (2010). Expressive robots in education - varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the 28th acm conference on human factors in computing systems (chi2010)*, atlanta.
- Salem, M., Kopp, S., Wachsmuth, I., Rohlfling, K., & Joublin, F. (2012). Generation and evaluation of communicative robot gesture. *International Journal of Social Robotics*, 4, 201-217.

- Scheutz, M., Schemerhorn, P., Kramer, J., & Anderson, D. (2007). First steps toward natural human-like HRI. *Auton Robot*, *22*, 411-423.
- Schulte, J., Rosenberg, C., & Thrun, S. (1999). Spontaneous, short-term interaction with mobile robots. In *Proceedings of the International Conference on Robotics and Automation*.
- Severinson-Eklundh, K., Green, A., & Hüttenrauch, H. (2003). Social and collaborative aspects of interaction with a service robot. *Robotics and Autonomous Systems*, *42*, 223-234.
- Shahid, S., Krahmer, E., & Swerts, M. (2011). Child-robot interaction: Playing alone or together? In *CHI 2011, May 7-12, 2011, Vancouver, BC, Canada*.
- Stifter, C. A., & Fox, N. A. (1987). Preschool children's ability to identify and label emotions. *Journal of Nonverbal Behavior*, *11*(1), 43-54.
- Tesser, F., Zovato, E., Nicolao, M., & Cosi, P. (2010). Two vocoder techniques for neutral to emotional timbre conversion. In *7th speech synthesis workshop (SSW)*.
- Van Dam, I. (2013). *Masters thesis*. (Unpublished)
- Wundt, W. M. (1897). *Outlines of Psychology*. York University, Toronto, Ontario.
- Yarczower, M., & Daruns, L. (1982). Social inhibition of spontaneous facial expressions in children. *Journal of Personality and Social Psychology*, *43*(4), 831-837.
- Zalm, A. Van der. (2011). *Help, i need some body!* Unpublished master's thesis, Universiteit Utrecht.

Appendix 1: Use Cases and Requirements

Requirements:

1. Can receive information about user

(a) Monitor Emotions User (R001)

- i. The robot should be able to monitor the emotions of the user while interacting to give appropriate emotional responses.

(b) Model Extroversion User (R002)

- i. The robot should be able to monitor the extroversion of the child, if this corresponds correctly to its own extroversion level.

2. Has internal model of users current state

(a) Arousal model of User (R003)

- i. The robot should have an internal representation of the arousal of the user.

(b) Valence model of User (R004)

- i. The robot should have an internal model of the valence of the emotion of the user to reason with.

3. Show Emotion

(a) Body poses for emotion (R005)

- i. The robot should know appropriate body movement and postures for certain combinations of arousal/valence.

(b) Convey emotion through voice (R006)

- i. The robot should be able to convey emotion through voice, for instance by adapting pitch and speech rate to the emotional state.

(c) Arousal (R007)

- i. The robot should have an internal representation of its arousal, depending on the situation and the emotional state of the child.
- (d) Valence (R008)
 - i. The robot should have an internal representation of its valence, depending on the situation and the emotional state of the child.
- (e) Knows which occurrences are emotionally relevant (R009)
 - i. The robot has a list of occurrences during the task being performed which tell which emotion is applicable to which occurrence. An example is winning (happy) or losing (sad) a game, or receiving a positive (happy) or negative (sad) answer when asking the child if it wants to play again. This list includes for every occurrence the corresponding arousal and valence value.
- (f) Emotional head position (R010)
 - i. The robot should be able to adapt the position of its head to its emotional state.
- (g) Emotional trunk position (R011)
 - i. The robot should be able to adapt the position of its trunk to its emotional state.

4. Is capable of gesturing in a natural way

- (a) Can adapt gesture size to internal state (R012)
 - i. The robot can adapt its gesture size to its emotional state and extroversion level.
- (b) Knows the appropriate movements for dialogue (R013)
 - i. The robot knows which gestures are fitting for which part of the dialogue.

Use-cases:

UC001	Robot mimics emotion child
<i>Pre-condition</i>	The robot had input about the valence and arousal of the child.
<i>Post-condition</i>	The child feels understood by the robot and likes it.
<i>Action sequence main</i>	<ol style="list-style-type: none"> 1. Woz input about the arousal and valence of the child. 2. The robot adapts its own arousal and valence towards that of the child. For instance, the input is A10 and V15, the robot had A9 and V19, the robot will adapt to A10 and V17. 3. No special occurrence linked with a pose happened, one body pose for emotion is not chosen. Trunk position is adapted to valence values. Gesture size, head position, speech volume, fundamental frequency of speech and speech rate are adapted to arousal value. 4. The robot expresses its emotional state subtly through the channels mentioned above.
<i>Actors</i>	Child, Robot
<i>Requirements</i>	R001, R003, R004, R006, R007, R008, R010, R011, R012

UC002	Valence child drops too low
<i>Pre-condition</i>	Valence of the child's emotion drops below a certain threshold.
<i>Post-condition</i>	The valence of the child goes up again.
<i>Action sequence main</i>	<ol style="list-style-type: none"> 1. Woz input about the valence of the child drops below threshold, say it is 5 out of 20. 2. Since the valence of the child is so low, the valence of the robot will not automatically follow the valence of the child. Instead, the valence of the robot will go towards a certain value, say 12. This value is chosen in such a way that the robot will not be too sad, but also not too happy. (We do not want the child to think the robot is happy its sad) If the robot had a valence of 6 before it will become 9, if its valence was 13 it will become 12, etc. 3. No special occurrence linked with a pose happened, one body pose for emotion is not chosen. The trunk position of the robot moves towards the fitting position.
<i>Actors</i>	Child, Robot
<i>Requirements</i>	R001, R004, R008, R011

UC003	Arousal child becomes too high
<i>Pre-condition</i>	Arousal of the child's emotion surpasses a certain threshold.
<i>Post-condition</i>	Arousal of the child drops a bit so it can concentrate better.
<i>Action sequence main</i>	<ol style="list-style-type: none"> 1. Woz input about the arousal of the child surpasses a threshold, say 16 out of 20. 2. Since the arousal of the child is so high, the arousal of the robot wont necessarily follow in the same direction, but instead go towards a certain value, say 6. If the robot had an arousal of 8 it will become 7, if it was 14 it will become 10, etc. 3. No special occurrence linked with a pose happened, one body pose for emotion is not chosen. Speech volume, Fundamental frequency of the voice, speech rate, gesture size and head position will be adapted to the current arousal value of the robot. This will make the robot give an overall calmer impression.
<i>Actors</i>	Child, Robot
<i>Requirements</i>	R001, R003, R006, R007, R010, R012

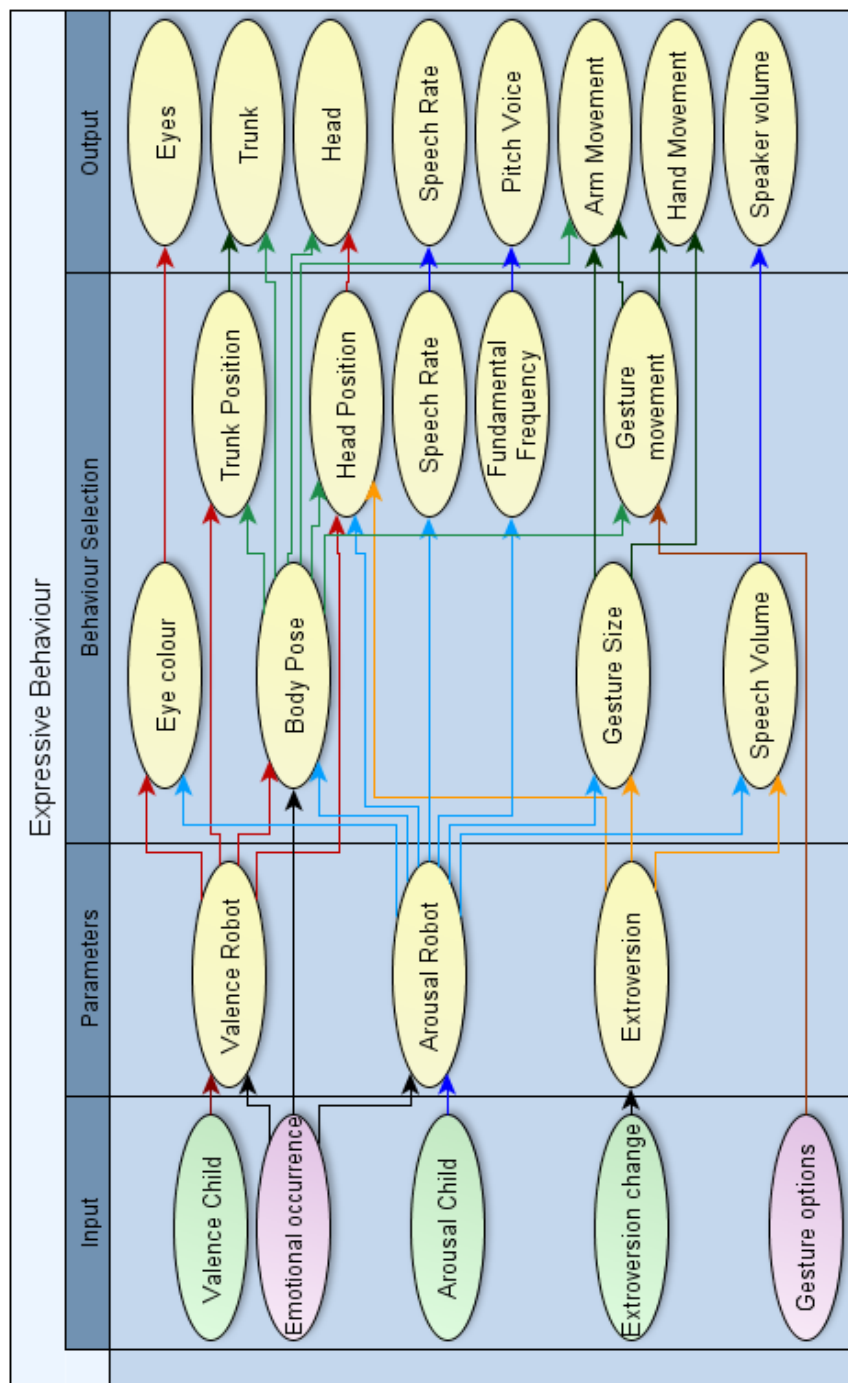
UC004	Robot reacts to emotional relevant occurrence
<i>Pre-condition</i>	Something happens which is correlated with a specific strong emotion.
<i>Post-condition</i>	Child understands emotion of the robot.
<i>Action sequence main</i>	<ol style="list-style-type: none"> 1. Something happens during the activity which is linked with a specific strong emotion for the robot. An example is if the robot loses a game, or if the child reacts positively to the question if it wishes to play again. 2. The arousal and the valence of the robot are adapted to the specific emotion. 3. A body pose belonging to the specific emotion is chosen, randomly if multiple poses for this emotion exist. 4. A sign is given to the modules for head movement, trunk position and gesture movement that they do not need to generate anything, since the body pose needs to be expressed first. 5. The robot displays the body pose
<i>Actors</i>	Child, Robot
<i>Requirements</i>	R005, R009

UC005	Robot is excited while talking
<i>Pre-condition</i>	The arousal and valence of the robot are high.
<i>Post-condition</i>	The child recognizes that the robot has different emotional states and attributes feelings to the robot.
<i>Action sequence main</i>	<ol style="list-style-type: none"> 1. The robots arousal and valence are high. 2. The robots speech volume, fundamental frequency in speech, speech rate and gesture size are high, the head up due to a high arousal. 3. The robots trunk becomes stretched due to a high valence. 4. The robot shows these aspects while speaking, so it speaks quicker and louder than normal, gestures big and has an upright head and stretched trunk.
<i>Actors</i>	Child, Robot
<i>Requirements</i>	R006, R007, R008, R010, R011, R012, R013

UC006	Robot talks while calm
<i>Pre-condition</i>	The robot has average valence and low arousal.
<i>Post-condition</i>	The child recognizes that the robot has different emotional states and attributes feelings to the robot.
<i>Action sequence main</i>	<ol style="list-style-type: none"> 1. The robots arousal is low, valence is average. 2. The robots speech volume, fundamental frequency in speech, speech rate and gesture size are low, the head low due to a low arousal. 3. The robots trunk is in average position due to an average valence. 4. The robot shows these aspects while speaking, so it speaks somewhat slower and softer than normal, gestures not to largely and has an somewhat lowered head and normal trunk position.
<i>Actors</i>	Child, Robot
<i>Requirements</i>	R006, R007, R008, R010, R011, R012, R013






















UC007	Robot has extroversion level which can be adapted to the child
<i>Pre-condition</i>	We know if the child is extrovert or introvert. Robot has extroversion level based on extroversion level of the child.
<i>Post-condition</i>	The robot expresses itself in a way fitting to the extroversion of the child, which makes the child like the robot more.
<i>Action sequence main</i>	<ol style="list-style-type: none"> 1. The robot has a scale for extroversion and its value is adapted to the extroversion of the child. 2. If extroversion is higher than average, speech volume, gesture size and head position will have a higher mean value. If extroversion is lower than average, speech volume, gesture size and head position will have a lower mean value. 3. There are upper and lower bounds for speech volume, gesture size and head position to ensure understandability and maintain eye contact with the child. 4. If the experimenter notices that the child behaves more/less extrovert than appeared from the test, the Woz can update the extroversion level of the robot in the right direction.
<i>Actors</i>	Child, Robot
<i>Requirements</i>	R002, R006, R010, R012

Appendix 2: Expressive Behaviour



Appendix 3: Eye colours

Eye colour of the Nao robot depending on Arousal

Arousal	R	G	B	Colour:
1	255	0	0	
0,9	255	51	0	
0,8	255	102	0	
0,7	255	153	0	
0,6	255	204	0	
0,5	255	255	0	
0,4	204	255	0	
0,3	153	255	0	
0,2	102	255	0	
0,1	51	255	0	
0	0	255	0	
-0,1	0	255	51	
-0,2	0	255	102	
-0,3	0	255	153	
-0,4	0	255	204	
-0,5	0	255	255	
-0,6	0	204	255	
-0,7	0	153	255	
-0,8	0	102	255	
-0,9	0	51	255	
-1	0	0	255	

Appendix 4: Use Case Quiz

	Ar. Ch. -1 -1	Val. Ch. -1 -1	Dialogue robot	Task state	Speech child
1	0,4	0,5	Hallo, ik ben Hugo, hoe heet jij?	Neutral	
2					Charlie
3	0,5	0,5	Ik wil graag de quiz met je spelen. Weet je hoe het moet?	Neutral	
4					ja
5	0,6	0,6	Jij mag beginnen. Welk onderwerp kies je?	Neutral	
6					de schijf van 5
7	0,5	0,6	<i>Stelt vraag en geeft mogelijke antwoorden</i>	Neutral	
8					Ik denk antwoord 2
9	0,4	0,3	Helaas, dat is fout. Probeer het nog een keer.	Neutral	
10					Antwoord 3
11	0,2	-0,1	Helaas, fout. Ik ben aan de beurt. Ik kies het onderwerp energiebalans	Neutral	
12					<i>Stelt vraag & geeft mogelijke antwoorden</i>
13	0	-0,4	Is het antwoord 4?	Neutral	
14					Dat is goed
15	-0,1	-0,3	Hoera! Ik sta nu voor. Jij bent aan de beurt. Welk onderwerp kies je?	Happy	
16					energiebalans
17	0	-0,2	<i>Stelt vraag en geeft mogelijke antwoorden</i>	Neutral	
18					3?
19	-0,1	-0,3	Helaas, fout. Probeer het nog een keer.	Neutral	
20					1?
21	-0,1	-0,5	Helaas, dat is fout. Ik sta nog steeds voor. Ik ben aan de beurt, ik kies de schijf van 5.	Neutral	
22					<i>Stelt vraag & geeft mogelijke antwoorden</i>
23	-0,2	-0,75	Antwoord 1?	Neutral	
24					dat is goed
25	0	-0,4	Hoera! Ik sta nog steeds voor. Het is alweer tijd om te stoppen. Vond je het leuk?	Happy	
26					nee
27	-0,1	-0,2	Wat jammer, ik vond het wel leuk.	Sad	

Input values in the use case for a quiz.

Ar. Ch = Arousal Child, Val. Ch = Valence Child, Gaze ch = Gaze direction child

Extroversion 0-100	Arousal R -1 -1	Valence R -1 -1	Eye colour	SV 45-65	FF 1-1.30	SR 85-105	Head -10 -16	Trunk 6 -18	Gesture M	Gesture S 0-100
40	0.2	0.5		57	1.18	97	5.6	13.5	Iconic	50
40	0.4	0.5		59	1.21	99	8.2	13.5	Iconic	55
40	0.5	0.6		60	1.23	100	9.5	14.0	Beat	58
40	0.5	0.6		60	1.23	100	9.5	14.0	Iconic	58
40	0.4	0.4		59	1.21	99	8.2	13.0	Beat	55
40	0.3	0.1		58	1.19	98	6.9	11.5	Beat	53
40	0.1	-0.2		56	1.17	96	4.3	10.0	Deictic	48
40	0.5	0.4		60	1.23	100	Happy	Happy	Happy	Happy
40	0.2	0.1		57	1.18	97	5.6	11.5	Iconic	50
40	0	-0.1		55	1.15	95	3	10.5	Beat	45
40	-0.1	-0.3		54	1.14	94	1.7	9.5	Beat	43
40	-0.2	-0.1		53	1.12	93	0.4	10.5	Iconic	40
40	0.4	0.4		59	1.21	99	Happy	Happy	Happy	Happy
40	-0.3	-0.4		52	1.11	92	Sad	Sad	Sad	Sad

Behaviour parameters and selection.

SV = Speech Volume, FF = Fundamental Frequency voice, SR = Speech Rate, Head = Head position, Trunk = Trunk position, Gesture M = Gesture Movement, Gesture S = Gesture size

Appendix 5: GOAL code

This appendix will describe my GOAL code in detail. The first section will give a short introduction in prolog, after which the different components of the program will be clarified.

Prolog

Because GOAL is a prolog-based language, its structure is quite different from most programs written in imperative languages. Before my program can be fully described, it is necessary to give some more information about the language. Prolog is a programming language based in logic, a prolog program consists of facts and rules with either constants or variables. The program can be used by asking it questions. A simple example of a prolog program can be found in below.

1. `woman(queen).`
2. `woman(princess).`
3. `man(king).`
4. `parent(queen,princess).`
5. `parent(king,princess).`
6. `mother(X,Y) :- woman(X), parent(X,Y).`
7. `daughter(X,Y) :- woman(X), parent(Y,X).`

This program tells us several things. First, it states that *queen* and *princess* are women (1 & 2) and that *king* is a man (3). Secondly, it tells us that *queen* and *king* are parents of *princess* (4 & 5). These are all facts, with *queen*, *princess* and *king* as constants. At line 6 & 7, we see two rules. 6 states that X is a mother of Y when X is a woman and X is a parent of Y. 7 specifies that X is a daughter of Y when X is a woman and Y is a parent of X. These two are rules, with X and Y as variables. Everything within brackets with a capital is a variable, without capital a

constant. If we were to ask the program in Table X the question $mother(queen, Y)$, prolog will try to give an instantiation for the variable Y which complies with the rules and facts. This program would return $Y = princess$. If we were to ask $mother(king, Y)$, prolog would return no , because $king$ is not a mother according to the program.

GOAL is prolog-based, which means that it works with rules and facts, but it has some additional functions. These functions will be explained in the next sections when they occur in my code.

Knowledge Base

Random number generator so chance can be calculated

random(X) is a prolog function which will give a random number between 0 and X. This function will therefore initialize Chance to a random number between 0 and 100.

chanceNumber(Chance) :- Chance is random(100).

Definitions which state when time-related actions need to be performed

TimePassed is the time which has passed between the old time (Y, saved in belief base) and the current time (X, from prolog function which gets the current time).

timeSinceUpdate(TimePassed) :- get_time(X), oldTime(Y), TimePassed is (X - Y).

States when an update is needed, namely when at least 30 seconds have passed.

updateNodig :- timeSinceUpdate(TimePassed), TimePassed >29.

Defines how much time has passed since the robot might have looked down. This is the difference between Y (last time the robot mightve looked down, from belief base) and X (current time from prolog function)

timeSinceLookDown(TimePassed) :- get_time(X), oldTimeHead(Y), TimePassed is (X - Y).

Defines when the robot might look down, namely when at least 30 seconds have passed.

maylookdown :- timeSinceLookDown(TimePassed), TimePassed >29.

Determines if the robot looks down

If the randomly generated number Chance is below 100 - the Extroversion of the robot, the robot will look down. For example, if the Extroversion is 75%, the robot will look down if Chance is below 25. This means that the robot has a 25% chance to look down, since Chance is a random number between 0 and 100.

lookdown:- extroversion(Extroversion), chanceNumber(Chance), Chance <(100-Extroversion).

Define when arousal and valence are too high or low
<i>States that arousal of the child is too low when it is below -0.95</i>
lowarousal :- childArousal(ChildArousal), ChildArousal <-0.95.
<i>States that arousal of the child is too high when it is above 0.9</i>
higharousal :- childArousal(ChildArousal), ChildArousal >0.9.
<i>States that valence of the child is too low when it is below -0.75</i>
lowvalence :- childValence(ChildValence), ChildValence <-0.75.

Calculates what the arousal and valence are in percentages
<i>ArousalPercentage and ValencePercentage represent the value of arousal and valence in percentage. This is useful because some other formulas need percentages.</i>
arousalPercentage(ArousalPercentage) :- arousal(Arousal), ArousalPercentage is (((Arousal + 1) / 2) * 100).
valencePercentage(ValencePercentage) :- valence(Valence), ValencePercentage is (((Valence + 1) / 2) * 100).

Formulas for prosodic values
<i>Speaker volume is determined by arousal and extroversion. This formula will give SpeechVolume a round value between SVMin and SVMax with a linear relationship with both Arousal and Extroversion.</i>
speechVolume(SpeechVolume) :- arousal(Arousal), extroversion(Extroversion), svMinimum(SVMin), svMaximum(SVMax), SpeechVolume is round((SVMin + (((Arousal+1) + (Extroversion / 50)) / 4) * (SVMax - SVMin))).
<i>Speech rate is determined by arousal. This formula will give SpeechRate a round value between SRMin and SRMax, linearly related to Arousal.</i>
speechRate(SpeechRate) :- arousal(Arousal), srMinimum(SRMin), srMaximum(SRMax), SpeechRate is round((SRMin + (((Arousal + 1) / 2) * (SRMax - SRMin))).
<i>Fundamental frequency is determined by arousal. This formula will give FundFrequency a value between FFMin and FFMax, linearly related to Arousal.</i>
fundFrequency(FundFrequency) :- arousal(Arousal), ffMinimum(FFMin), ffMaximum(FFMax), FundFrequency is (FFMin + (((Arousal + 1) / 2) * (FFMax - FFMin))).

Formulas for body positions
<i>Gesture size is determined by arousal and extroversion. This formula will give GestureSize a value between 0 and 100, linearly related to arousal and extroversion.</i>
gestureSize(GestureSize) :- arousal(Arousal), extroversion(Extroversion), GestureSize is (((((Arousal + 1) / 2) * 100) + Extroversion) / 2).
<i>Trunk position is determined by valence. This formula will give TrunkPosition a value between TrunkMin and TrunkMax, linearly related to Valence.</i>
trunk(TrunkPosition) :- valence(Valence), trunkMaximum(TrunkMax), trunkMinimum(TrunkMin), TrunkPosition is (TrunkMin + (((Valence + 1) / 2) * (TrunkMax - TrunkMin))).
<i>Head position is determined by valence and arousal. This formula will give HeadPosition a value between HeadMin and HeadMax, linearly related to valence and arousal.</i>
head(HeadPosition) :- arousalPercentage(ArousalPercentage), valencePercentage(ValencePercentage), headMaximum(HeadMax), headMinimum(HeadMin), HeadPosition is (HeadMax - (((ArousalPercentage + ValencePercentage) / 200) * (HeadMax - HeadMin))).

Formulas to calculate the RGB values for the eyes
<i>Eye values based on Arousal only. For instance, when Arousal is 0, eyeRed(0), eyeGreen(255) and eyeBlue(0) will hold.</i>
eyeRed(0) :- arousal(Arousal), Arousal <0.05, Arousal >-1.
eyeRed(51) :- arousal(Arousal), Arousal <0.15, Arousal >= 0.05.
eyeRed(102) :- arousal(Arousal), Arousal <0.25, Arousal >= 0.15.
eyeRed(153) :- arousal(Arousal), Arousal <0.35, Arousal >= 0.25.
eyeRed(204) :- arousal(Arousal), Arousal <0.45, Arousal >= 0.35.
eyeRed(255) :- arousal(Arousal), Arousal =<1, Arousal >= 0.45.
eyeGreen(0) :- arousal(Arousal), ((Arousal <-0.95, Arousal >= -1) ; (Arousal =<1, Arousal >= 0.95)).
eyeGreen(51) :- arousal(Arousal), ((Arousal <-0.85, Arousal >= -0.95) ; (Arousal <0.95, Arousal >= 0.85)).
eyeGreen(102) :- arousal(Arousal), ((Arousal <-0.75, Arousal >= -0.85) ; (Arousal <0.85, Arousal >= 0.75)).
eyeGreen(153) :- arousal(Arousal), ((Arousal <-0.65, Arousal >= -0.75) ; (Arousal <0.75, Arousal >= 0.65)).
eyeGreen(204) :- arousal(Arousal), ((Arousal <-0.55, Arousal >= -0.65) ; (Arousal <0.65, Arousal >= 0.55)).
eyeGreen(255) :- arousal(Arousal), Arousal <0.55, Arousal >= -0.55.
eyeBlue(255) :- arousal(Arousal), Arousal <-0.45, Arousal >= -1.
eyeBlue(204) :- arousal(Arousal), Arousal <-0.35, Arousal >= -0.45.
eyeBlue(153) :- arousal(Arousal), Arousal <-0.25, Arousal >= -0.35.
eyeBlue(102) :- arousal(Arousal), Arousal <-0.15, Arousal >= -0.25.
eyeBlue(51) :- arousal(Arousal), Arousal <-0.05, Arousal >= -0.15.
eyeBlue(0) :- arousal(Arousal), Arousal =<1, Arousal >= -0.05.
<i>Formulas which calculate the eye values based on the value for arousal only and the Valence. When valence is 0, eye values will stay the same. When it is below zero, the second formula holds, when it is above 0 the third.</i>
eyeValence(Xold, Xnew) :- valence(Valence), Valence = 0, Xnew = Xold.
eyeValence(Xold, Xnew) :- valence(Valence), (Valence <0), valencePercentage(ValencePercentage), Xnew is (((Xold * ((50 - ValencePercentage) * 2)) / 100)).
eyeValence(Xold, Xnew) :- valence(Valence), (Valence >0), valencePercentage(ValencePercentage), Xnew is (Xold + (((255 - Xold) * ((ValencePercentage - 50) * 2)) / 100)).

Formulas to calculate new arousal and valence values
<i>Definitions for how the current arousal (ArousalOld) is adapted towards a certain value (TowardsValue). Three formulas, which one is used depends on whether the arousal needs to be raised or lowered.</i>
adapta(ArousalOld,TowardsValue,ArousalNew) :- Verschil is ArousalOld - TowardsValue, Verschil >0.05, ArousalNew is ((ArousalOld - Verschil)/2).
adapta(ArousalOld,TowardsValue,ArousalNew) :- Verschil is TowardsValue - ArousalOld, Verschil >0.05, ArousalNew is (ArousalOld + Verschil/2).
adapta(ArousalOld,TowardsValue,ArousalNew) :- Verschil is TowardsValue - ArousalOld, -0.05 <Verschil, Verschil <0.05, ArousalNew = TowardsValue.
<i>Definitions for how the current valence (ValenceOld) is adapted towards a certain value (TowardsValue). Three formulas, which one is used depends on whether the valence needs to be raised or lowered.</i>
adaptv(ValenceOld,TowardsValue,ValenceNew) :- Verschil is ValenceOld - TowardsValue, Verschil >0.05, ValenceNew is (ValenceOld - Verschil/2).
adaptv(ValenceOld,TowardsValue,ValenceNew) :- Verschil is TowardsValue - ValenceOld, Verschil >0.05, ValenceNew is (ValenceOld + Verschil/2).
adaptv(ValenceOld,TowardsValue,ValenceNew) :- Verschil is TowardsValue - ValenceOld, -0.05 <Verschil, Verschil <0.05, ValenceNew = TowardsValue.

Formulas to choose a gesture
<i>Determine the % chance of Iconic, Beat, Deictic and Metaphoric gesture, given that the total of initial chances might not be 100. I stands for: initial % Chance of Iconic gesture B stands for: initial % Chance of Beat gesture D stands for: initial % Chance of Deictic gesture M stands for: initial % Chance of Metaphoric gesture</i>
chanceIconic(Chance) :- iP(I), bP(B), dP(D), mP(M), Chance is $((100 * I) / (I + B + D + M))$.
chanceBeat(Chance) :- iP(I), bP(B), dP(D), mP(M), Chance is $((100 * B) / (I + B + D + M))$.
chanceDeictic(Chance) :- iP(I), bP(B), dP(D), mP(M), Chance is $((100 * D) / (I + B + D + M))$.
chanceMetaphoric(Chance) :- iP(I), bP(B), dP(D), mP(M), Chance is $((100 * M) / (I + B + D + M))$.
<i>Decides which gesture kind is actually chosen with a random number Chance between 0 and 100. For example, an iconic gesture is chosen if Chance is below the chance for an Iconic gesture. A beat gesture is chosen if Chance is between the chance for an iconic gesture and the sum of the chances for an iconic and a beat gesture, etc.</i>
chooseGesture(iconic) :- chanceIconic(IconicChance), chanceNumber(Chance), IconicChance >0, X is IconicChance + 1, Chance <X.
chooseGesture(beat) :- chanceIconic(IconicChance), chanceBeat(BeatChance), chanceNumber(Chance), BeatChance >0, X is IconicChance + BeatChance + 1, Y is IconicChance, Chance <X, Chance >Y.
chooseGesture(deictic) :- chanceIconic(IconicChance), chanceBeat(BeatChance), chanceDeictic(DeicticChance), chanceNumber(Chance), DeicticChance >0, X is IconicChance + BeatChance + DeicticChance + 1, Y is IconicChance + BeatChance, Chance <X, Chance >Y.
chooseGesture(metaphoric) :- chanceIconic(IconicChance), chanceBeat(BeatChance), chanceDeictic(DeicticChance), chanceMetaphoric(MetaphoricChance), chanceNumber(Chance), MetaphoricChance >0, Y is IconicChance + BeatChance + DeicticChance, Chance >Y.
<i>Choose which gesture is acutally performed. Choices is a list with all possible gestures, Number the one chosen. It is chosen by randomly selecting the index number of the gesture chosen.</i>
chooseGestureNumber(Choices,Number) :- length(Choices, Length), random(0, Length, Index), nth0(Index, Choices, Number). chooseGestureNumber([], _) :- !, fail.

Belief Base

Start

This belief indicates that the program is in its first cycle. The first cycle, this belief will be removed and in its stead, the current time will be added to the belief base.

start.

Emotion and personality

Initial beliefs about the robot's own internal state, its own extroversion, arousal and valence.

extroversion(50).

arousal(0).

valence(0.5).

Initial beliefs about the valence and arousal of the child

childArousal(0).

childValence(0).

Behaviour parameters
<i>Gives the current value of the robot's speech volume. Initialized to 0. Every time the speech volume is updated, so is this belief.</i>
speechVolumeOld(0).
<i>Gives the current value of the robot's speech rate. Initialized to 0. Every time the speech rate is updated, so is this belief.</i>
speechRateOld(0).
<i>Gives the current value of the robot's fundamental frequency. Initialized to 0. Every time the fundamental frequency is updated, so is this belief.</i>
fundFrequencyOld(0).
<i>Gives the current value of the robot's head position. Initialized to 0. Every time the head position is updated, so is this belief.</i>
headOld(0).
<i>Gives the current value of the robot's trunk position. Initialized to 0. Every time the trunk position is updated, so is this belief.</i>
trunkOld(0).
<i>Gives the current value of the robot's eye colour. Initialized to 0,255,0. Every time the eye colour is updated, so is this belief.</i>
eyeOld(0,255,0).

Minimum and Maximum values for behaviour
<i>Initial minimum and maximum values for trunk position</i>
trunkMinimum(8).
trunkMaximum(14).
<i>Initial minimum and maximum values for head position</i>
headMinimum(-10).
headMaximum(16).
<i>Initial minimum and maximum values for Speech Volume</i>
svMinimum(45).
svMaximum(65).
<i>Initial minimum and maximum values for Speech Rate</i>
srMinimum(85).
srMaximum(105).
<i>Initial minimum and maximum values for Fundamental Fequency</i>
ffMinimum(1).
ffMaximum(4).

Action Base

Speech actions
<i>Send the new value for SpeechVolume to the WoOz.</i>
adaptSpeechVolume(SpeechVolume) {
pre{true}
post{true}
}
<i>Send the new value for SpeechRate to the WoOz.</i>
adaptSpeechRate(SpeechRate) {
pre{true}
post{true}
}
<i>Send the new value for FundFrequency to the WoOz.</i>
adaptFundFrequency(FundFrequency) {
pre{true}
post{true}
}

Body actions
<i>Send the new value for GestureSize to the WoOz.</i>
adaptGestureSize(GestureSize) {
pre{true}
post{true}
}
<i>Send the new value for TrunkPosition to the WoOz.</i>
adapttrunk(TrunkPosition) {
pre{true}
post{true}
}
<i>Send the new value for HeadPosition to the WoOz.</i>
adapthead(HeadPosition) {
pre{true}
post{true}
}

Actie om de R, G en B waarde door te geven
<i>Send the new eye values for R, G and B to the WoOz.</i>
adapteye(R,G,B) {
pre{true}
post{true}
}

Look down
<i>Tell the WoOz that the robot needs to look down</i>
lookdownaction{
pre{true}
post{true}
}

Gesture
<i>Send ID (Number) of the chosen gesture back to the WoOz, along with the gesture size.</i>
sendGesture(GestureNumber, GestureSize){
pre{true}
post{true}
}

Program Main

Rules for timely updates
<i>If an update of Valence and Arousal is needed, indicate in the belief base that this is the case and save the current time for future reference.</i>
if bel(updateNodig, get_time(X), oudtijd(Y))
then delete(oudtijd(Y)) + insert(oudtijd(X)) + insert(notmimickeda) + insert(notmimickedv).
<i>If the robot mightve looked down, but decided not to, save the current time for future reference.</i>
if bel(maylookdown, not(lookdown), get_time(X), oldTimeHead(Y))
then delete(oldTimeHead(Y)) + insert(oldTimeHead(X)).
<i>If the robot needs to look down, perform the action which does so and save the current time for future reference.</i>
if bel(maylookdown, lookdown, get_time(X), oldTimeHead(Y))
then delete(oldTimeHead(Y)) + insert(oldTimeHead(X)) + lookdownaction.

Adapt Arousal
<i>If the arousal of the child is not too high or too low, adapt arousal robot towards arousal child and indicate that arousal is up to date.</i>
if bel(not(lowarousal), not(higharousal), childArousal(ChildArousal), arousal(ArousalOld), notmimickeda, adapta(ArousalOld,ChildArousal,ArousalNew))
then delete(notmimickeda) + delete(arousal(ArousalOld)) + insert(arousal(ArousalNew)).
<i>If the arousal of the child is too low, adapt arousal robot towards -0.7 and indicate that arousal is up to date.</i>
if bel(lowarousal, arousal(ArousalOld), notmimickeda, adapta(ArousalOld,-0.7,ArousalNew))
then delete(notmimickeda) + delete(arousal(ArousalOld)) + insert(arousal(ArousalNew)).
<i>If the arousal of the child is too high, adapt arousal robot towards -0.4 and indicate that arousal is up to date.</i>
if bel(higharousal, arousal(ArousalOld), notmimickeda, adapta(ArousalOld,-0.4,ArousalNew))
then delete(notmimickeda) + delete(arousal(ArousalOld)) + insert(arousal(ArousalNew)).
<i>If an emotionally relevant occurrence just took place and the arousal has already been adapted towards the child, adapt arousal towards the value corresponding to the occurrence.</i>
if bel(arousalOccurrence(ArousalOcc), arousal(ArousalOld), not(notmimickeda), adapta(ArousalOld,ArousalOcc,ArousalNew))
then delete(arousalOccurrence(ArousalOcc)) + delete(arousal(ArousalOld)) + insert(arousal(ArousalNew)).

Adapt Valence
<i>If the child's valence is not too low, adapt valence robot towards valence child and indicate that valence is up to date.</i>
if bel(not(lowvalence), childValence(ChildValence), valence(ValenceOld), notmimickedv, adaptv(ValenceOld,ChildValence,ValenceNew))
then delete(notmimickedv) + delete(valence(ValenceOld)) + insert(valence(ValenceNew)).
<i>If the valence of the child is too low, adapt valence robot towards 0 and indicate that valence is up to date.</i>
if bel(lowvalence, valence(ValenceOld), notmimickedv, adaptv(ValenceOld,0,ValenceNew))
then delete(notmimickedv) + delete(valence(ValenceOld)) + insert(valence(ValenceNew)).
<i>If an emotionally relevant occurrence just took place and the valence has already been adapted towards the child, adapt valence towards the value corresponding to the occurrence.</i>
if bel(valenceOccurrence(ValenceOcc), valence(ValenceOld), not(notmimickedv), adaptv(ValenceOld,ValenceOcc,ValenceNew))
then delete(valenceOccurrence(ValenceOcc)) + delete(valence(ValenceOld)) + insert(valence(ValenceNew)).

Adapt Speech
<i>If the current speech volume according to the knowledgebase is not the same as the last speech volume sent to the robot (according to the belief base), send the correct speech volume value to the robot and adapt the belief base to this value.</i>
if bel(speechVolumeOld(SpeechVolumeOld), speechVolume(SpeechVolume), not(SpeechVolumeOld = SpeechVolume))
then delete(speechVolumeOld(SpeechVolumeOld)) + insert(speechVolumeOld(SpeechVolume)) + adaptSpeechVolume(SpeechVolume).
<i>If the current speech rate according to the knowledgebase is not the same as the last speech rate sent to the robot (according to the belief base), send the correct speech rate value to the robot and adapt the belief base to this value.</i>
if bel(speechRateOld(SpeechRateOld), speechRate(SpeechRate), not(SpeechRateOld = SpeechRate))
then delete(speechRateOld(SpeechRateOld)) + insert(speechRateOld(SpeechRate)) + adapt- SpeechRate(SpeechRate).
<i>If the current fundamental frequency according to the knowledgebase is not the same as the last fundamental frequency sent to the robot (according to the belief base), send the correct fundamental frequency value to the robot and adapt the belief base to this value.</i>
if bel(fundFrequencyOld(FundFrequencyOld), fundFrequency(FundFrequency), not(FundFrequencyOld = FundFrequency))
then delete(fundFrequencyOld(FundFrequencyOld)) + insert(fundFrequencyOld(FundFrequency)) + adaptFundFrequency(FundFrequency).

Adapt Body
<i>If the current trunk position according to the knowledgebase is not the same as the last trunk position sent to the robot (according to the belief base), send the correct trunk position value to the robot and adapt the belief base to this value.</i>
if bel(trunkOld(TrunkOld), trunk(Trunk), not(TrunkOld = Trunk))
then delete(trunkOld(TrunkOld)) + insert(trunkOld(Trunk)) + adapttrunk(Trunk).
<i>If the current head position according to the knowledgebase is not the same as the last head position sent to the robot (according to the belief base), send the correct head position value to the robot and adapt the belief base to this value.</i>
if bel(headOld(HeadOld), head(Head), not(HeadOld = Head))
then delete(headOld(HeadOld)) + insert(headOld(Head)) + adapthead(Head).

Adapt Eyes

If the current eye colours according to the knowledgebase rules are not the same as the last eye colours sent to the robot (according to the belief base), send the correct eye colours value to the robot and adapt the belief base to this value.

```
if bel(eyeRed(RedA), eyeGreen(GreenA), eyeBlue(BlueA), eyeValence(RedA, RedNew), eyeValence(GreenA, GreenNew), eyeValence(BlueA, BlueNew), eyeOld(RedOld, GreenOld, BlueOld), (not(RedOld = RedNew) ; not(GreenOld = GreenNew) ; not(BlueOld = BlueNew)) )
```

```
then delete(eyeOld(RedOld, GreenOld, BlueOld)) + insert(eyeOld(RedNew, GreenNew, BlueNew)) + adapteye(RedNew, GreenNew, BlueNew).
```

Gesture movement
<i>If a gesture has to be chosen and an iconic gesture is chosen with the ID number Number, send this Gesture to the robot and delete beliefs about this gesture choice.</i>
if bel(selectGesture(NrIconic,NrBeat,NrDeictic,NrMetaphoric), chooseGesture(iconic), chooseGestureNumber(NrIconic, Number), iP(Iconic), bP(Beat), dP(Deictic), mP(Metaphoric), gestureSize(GestureSize))
then delete(selectGesture(NrIconic,NrBeat,NrDeictic,NrMetaphoric)) + delete(iP(Iconic)) + delete(bP(Beat)) + delete(dP(Deictic)) + delete(mP(Metaphoric)) + sendGesture(Number, GestureSize).
<i>If a gesture has to be chosen and an beat gesture is chosen with the ID number Number, send this Gesture to the robot and delete beliefs about this gesture choice.</i>
if bel(selectGesture(NrIconic,NrBeat,NrDeictic,NrMetaphoric), chooseGesture(beat), chooseGestureNumber(NrBeat, Number), iP(Iconic), bP(Beat), dP(Deictic), mP(Metaphoric), gestureSize(GestureSize))
then delete(selectGesture(NrIconic,NrBeat,NrDeictic,NrMetaphoric)) + delete(iP(Iconic)) + delete(bP(Beat)) + delete(dP(Deictic)) + delete(mP(Metaphoric)) + sendGesture(Number, GestureSize).
<i>If a gesture has to be chosen and an deictic gesture is chosen with the ID number Number, send this Gesture to the robot and delete beliefs about this gesture choice.</i>
if bel(selectGesture(NrIconic,NrBeat,NrDeictic,NrMetaphoric), chooseGesture(deictic), chooseGestureNumber(NrDeictic, Number), iP(Iconic), bP(Beat), dP(Deictic), mP(Metaphoric), gestureSize(GestureSize))
then delete(selectGesture(NrIconic,NrBeat,NrDeictic,NrMetaphoric)) + delete(iP(Iconic)) + delete(bP(Beat)) + delete(dP(Deictic)) + delete(mP(Metaphoric)) + sendGesture(Number, GestureSize).
<i>If a gesture has to be chosen and an metaphoric gesture is chosen with the ID number Number, send this Gesture to the robot and delete beliefs about this gesture choice.</i>
if bel(selectGesture(NrIconic,NrBeat,NrDeictic,NrMetaphoric), chooseGesture(metaphoric), chooseGestureNumber(NrMetaphoric, Number), iP(Iconic), bP(Beat), dP(Deictic), mP(Metaphoric), gestureSize(GestureSize))
then delete(selectGesture(NrIconic,NrBeat,NrDeictic,NrMetaphoric)) + delete(iP(Iconic)) + delete(bP(Beat)) + delete(dP(Deictic)) + delete(mP(Metaphoric)) + sendGesture(Number, GestureSize).

Event Module

Percept rules about the emotions of the child

If a percept about the arousal of the child enters the system, remove the old beliefs about the child's arousal, add the new arousal as belief and indicate that the robot's arousal is no longer up to date.

```
forall bel(percept(arousalChild(ChildArousalNew)), childArousal(ChildArousalOld))
```

```
do delete(childArousal(ChildArousalOld)) + insert(childArousal(ChildArousalNew)) + insert(notmimickedv).
```

If a percept about the valence of the child enters the system, remove the old beliefs about the child's valence, add the new valence as belief and indicate that the robot's valence is no longer up to date.

```
forall bel(percept(valenceChild(ChildValenceNew)), childValence(ChildValenceOld))
```

```
do delete(childValence(ChildValenceOld)) + insert(childValence(ChildValenceNew)) + insert(notmimickedv).
```

Percept rule about extroversion

If the extroversion level of the child changes, remove the old extroversion level from the belief base and add the new one.

```
forall bel(percept(extroversionlevel(ExtroversionNew)), extroversion(ExtroversionOld))
```

```
do delete(extroversion(ExtroversionOld)) + insert(extroversion(ExtroversionNew)).
```

Percept rules about emotional relevant occurrences

If an emotional relevant occurrence took place, add corresponding arousal & valence to belief base and indicate that the robot's arousal and valence are no longer up to date.

```
forall bel(percept(occurrence(ArousalOcc, ValenceOcc)))
```

```
do insert(arousalOccurrence(ArousalOcc)) + insert(valenceOccurrence(ValenceOcc)) + insert(notmimickedv) + insert(notmimickedv).
```


Percept rules about chances in Minimum/Maximum values behaviours
<i>If a percept about new Minimum or Maximum values for Trunk, Head, Speech Volume, Speech Rate or Fund. Frequency occurs, enter this into the belief base.</i>
forall bel(percept(trunkMinimum(TrunkMin)), trunkMinimum(TrunkMinOld))
do delete(trunkMinimum(TrunkMinOld)) + insert(trunkMinimum(TrunkMin)).
forall bel(percept(trunkMaximum(TrunkMax)), trunkMaximum(TrunkMaxOld))
do delete(trunkMaximum(TrunkMaxOld)) + insert(trunkMaximum(TrunkMax)).
forall bel(percept(headMinimum(HeadMin)), headMinimum(HeadMinOld))
do delete(headMinimum(HeadMinOld)) + insert(headMinimum(HeadMin)).
forall bel(percept(headMaximum(HeadMax)), headMaximum(HeadMaxOld))
do delete(headMaximum(HeadMaxOld)) + insert(headMaximum(HeadMax)).
forall bel(percept(svMinimum(SVMin)), svMinimum(SVMinOld))
do delete(svMinimum(SVMinOld)) + insert(svMinimum(SVMin)).
forall bel(percept(svMaximum(SVMax)), svMaximum(SVMaxOld))
do delete(svMaximum(SVMaxOld)) + insert(svMaximum(SVMax)).
forall bel(percept(srMinimum(SRMin)), srMinimum(SRMinOld))
do delete(srMinimum(SRMinOld)) + insert(srMinimum(SRMin)).
forall bel(percept(srMaximum(SRMax)), srMaximum(SRMaxOld))
do delete(srMaximum(SRMaxOld)) + insert(srMaximum(SRMax)).
forall bel(percept(ffMinimum(FFMin)), ffMinimum(FFMinOld))
do delete(ffMinimum(FFMinOld)) + insert(ffMinimum(FFMin)).
forall bel(percept(ffMaximum(FFMax)), ffMaximum(FFMaxOld))
do delete(ffMaximum(FFMaxOld)) + insert(ffMaximum(FFMax)).

Percept rules about tekst related gestures
<i>If a narrative gesture has to be chosen from NrIconic number of Iconic gestures, NrBeat number of beat gestures, NrDeictic number of deictic gestures and NrMetaphoric number of metaphoric gesture, insert these numbers into the belief base, along with the initial chances for each kind of gestures given that the text is narrative. (For instance, the initial chance of an iconic gesture is $iP(36.94)$ is 0.3694)</i>
forall bel(percept(gestureOptions(narrative,NrIconic,NrBeat,NrDeictic,NrMetaphoric)))
do insert(selectGesture(NrIconic,NrBeat,NrDeictic,NrMetaphoric)) + insert(iP(36.94)) + insert(bP(21.73)) + insert(dP(39.11)) + insert(mP(2.22)).
<i>If a extranarrative gesture has to be chosen from NrIconic number of Iconic gestures, NrBeat number of beat gestures, NrDeictic number of deictic gestures and NrMetaphoric number of metaphoric gesture, insert these numbers into the belief base, along with the initial chances for each kind of gestures given that the text is extranarrative. (For instance, the initial chance of an iconic gesture is $iP(17.26)$ is 0.1726)</i>
forall bel(percept(gestureOptions(extraNarrative,NrIconic,NrBeat,NrDeictic,NrMetaphoric)))
do insert(selectGesture(NrIconic,NrBeat,NrDeictic,NrMetaphoric)) + insert(iP(17.26)) + insert(bP(57.52)) + insert(dP(11.51)) + insert(mP(13.71)).
<i>If one of the Number lists for a gesture is only has a 0, this means no gestures of this kind exist, so the initial chance for this gesture becomes 0.</i>
if bel(selectGesture ([NI—],NB,ND,NM), NI = 0, iP(X), not(X = 0))
then delete(iP(X)) + insert(iP(0)).
if bel(selectGesture (NI,[NB—],ND,NM), NB = 0, bP(X), not(X = 0))
then delete(bP(X)) + insert(bP(0)).
if bel(selectGesture (NI,NB,[ND—],NM), ND = 0, dP(X), not(X = 0))
then delete(dP(X)) + insert(dP(0)).
if bel(selectGesture (NI,NB,ND,[NM—]), NM = 0, mP(X), not(X = 0))
then delete(mP(X)) + insert(mP(0)).
Initialize the time
<i>If this is the first cycle, save the current time so we can measure how much time has passed.</i>
if bel(start, get_time(Y))
then insert(oldTime(Y)) + insert(oldTimeHead(Y)) + delete(start).

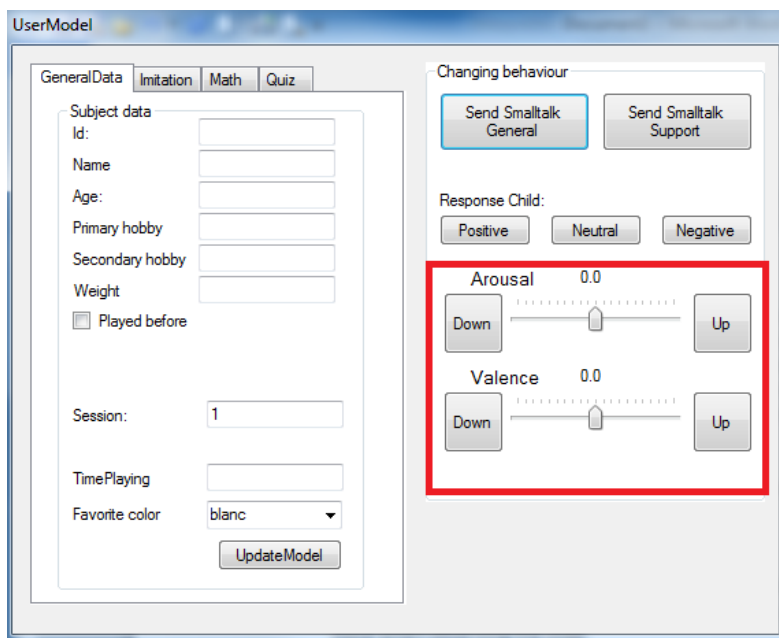
Appendix 6: Instructions

Arousal/Valence input

Instructions Method 1:

Your task is to interpret the emotional state of the child on the video. You will do this with the help of two parameters, namely Arousal and Valence. Arousal refers to how excited the child is, Valence to how positive the emotion is. For instance, boredom is an emotion with a very low arousal and a low valence, while happiness has a high arousal and valence. Both valence and arousal are represented on a scale from -1 to 1. This means that when arousal or valence is -1, it could not be any lower and when it is 1 it could not be any higher. A valence of 1 thus represents the most positive emotion, an arousal of 1 the most excited emotion, a valence of -1 the most negative emotion, arousal of -1 the least excited emotion.

You can interpret the emotion of the child using the following set-up:



The screenshot shows the 'UserModel' software interface. On the left, there is a 'GeneralData' tab with fields for 'Subject data' (Id, Name, Age, Primary hobby, Secondary hobby, Weight, Played before checkbox), 'Session' (1), 'TimePlaying', and 'Favorite color' (blanc). An 'UpdateModel' button is at the bottom. On the right, the 'Changing behaviour' section includes 'Send Smalltalk General' and 'Send Smalltalk Support' buttons, and 'Response Child' buttons for 'Positive', 'Neutral', and 'Negative'. The 'Arousal' and 'Valence' sliders are highlighted with a red box. Both sliders are currently set to 0.0, with 'Down' and 'Up' buttons on either side of the slider.

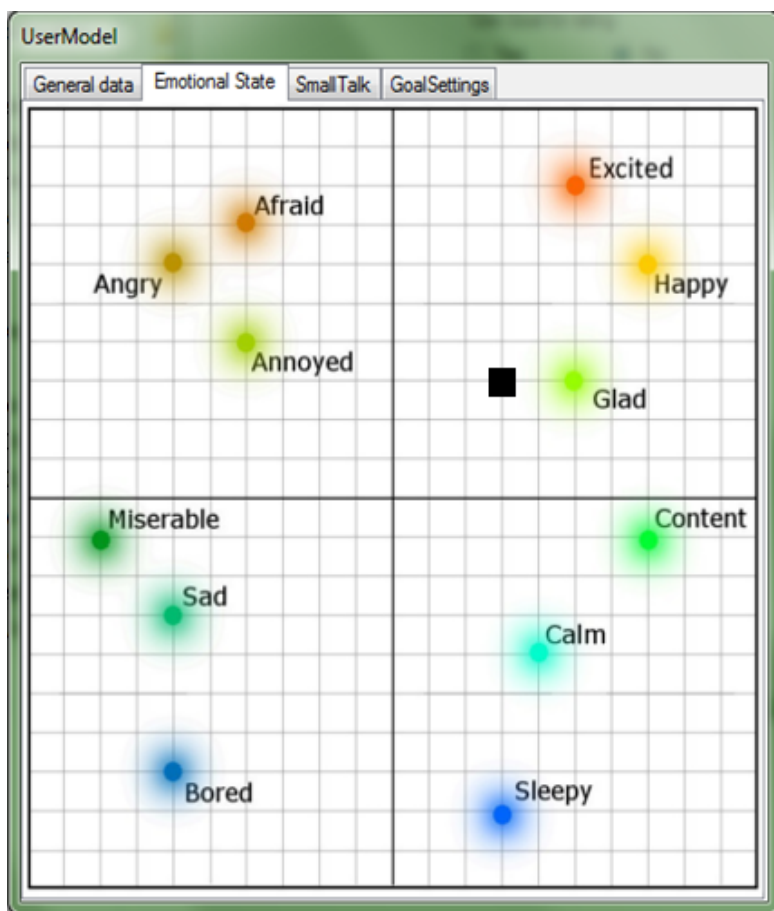
You will only need the part which is outlined in red, all the other fields can be ignored. You can adapt the valence and arousal by pressing the Up and Down buttons, or by dragging the slider. Pressing a button will always raise or lower the value by 0.1. You can press the buttons as often as you wish. When the video starts, you may start interpreting the emotional state of the child. Try to do this as precisely as possible.

Because interpreting something as arousal and valence is subjective, several guidelines have been established. During the experiment, you should try to stick to these guidelines as much as possible.

- The first guideline is a list of the behaviours mentioned in Table 6.1 and 6.2. Aside from the obvious facial expressions to interpret emotional state, these behaviours should be kept in mind when observing the child.
 - When you notice a rise in speech rate, arousal has probably gone up.
 - When you notice a rise in speech volume, arousal has probably gone up.
 - When you notice a rise in fundamental frequency, arousal has probably gone up.
 - When you notice an increase in gesture size, arousal has probably gone up.
 - When the head is higher than average, arousal and valence are probably above 0.
 - When the trunk is in a more stretched position than average, valence is probably above 0.
 - When the child smiles, valence is probably above 0 and goes up.
 - When the child laughs, valence is probably above 0.4 and goes up.
 - When the child frowns, valence is probably below 0 and goes down.
 - When the child cries, valence is probably below -0.4 and goes down.
 - When the child starts bouncing, arousal is probably above 0 and goes up.
 - When the child starts shrugging, arousal is probably below 0 and goes down.
- All of the before mentioned behaviours can also have opposite effects, examples listed below. Notice that the absence of certain behaviours (such as smiles) is also significant!
 - When you notice a decline in speech rate, arousal has probably gone down.
 - When head is lower than average, arousal and valence are probably below 0.
 - When the child stops smiling, valence probably goes down.
 - When the child stops shrugging, arousal probably goes up.
 - Etc.

- Arousal and Valence work on a scale from -1 to 1. This means that if they are both -1, the child cannot be any sadder and if they're both 1 the child cannot be happier. This should be taken into account when approaching these values. If Valence is already 0.9 the child should appear very happy, otherwise something might be wrong with the interpretation.
- The value of 0 for both Arousal and Valence represent the middle between the lowest and highest possible values. This means that Arousal 0 and Valence 0 do not necessarily represent the default emotional state of the child! It can very well be that the child is constantly calm and content, which would be better represented by, for instance, an Arousal of -0.4 and Valence of 0.3.

Instructions Method 2: The only part of the instructions which differed per method was the description of the input method and corresponding picture. In the above description, these are the Figure of the WoOz and the paragraph directly below it. For the second input method, these were as follows:

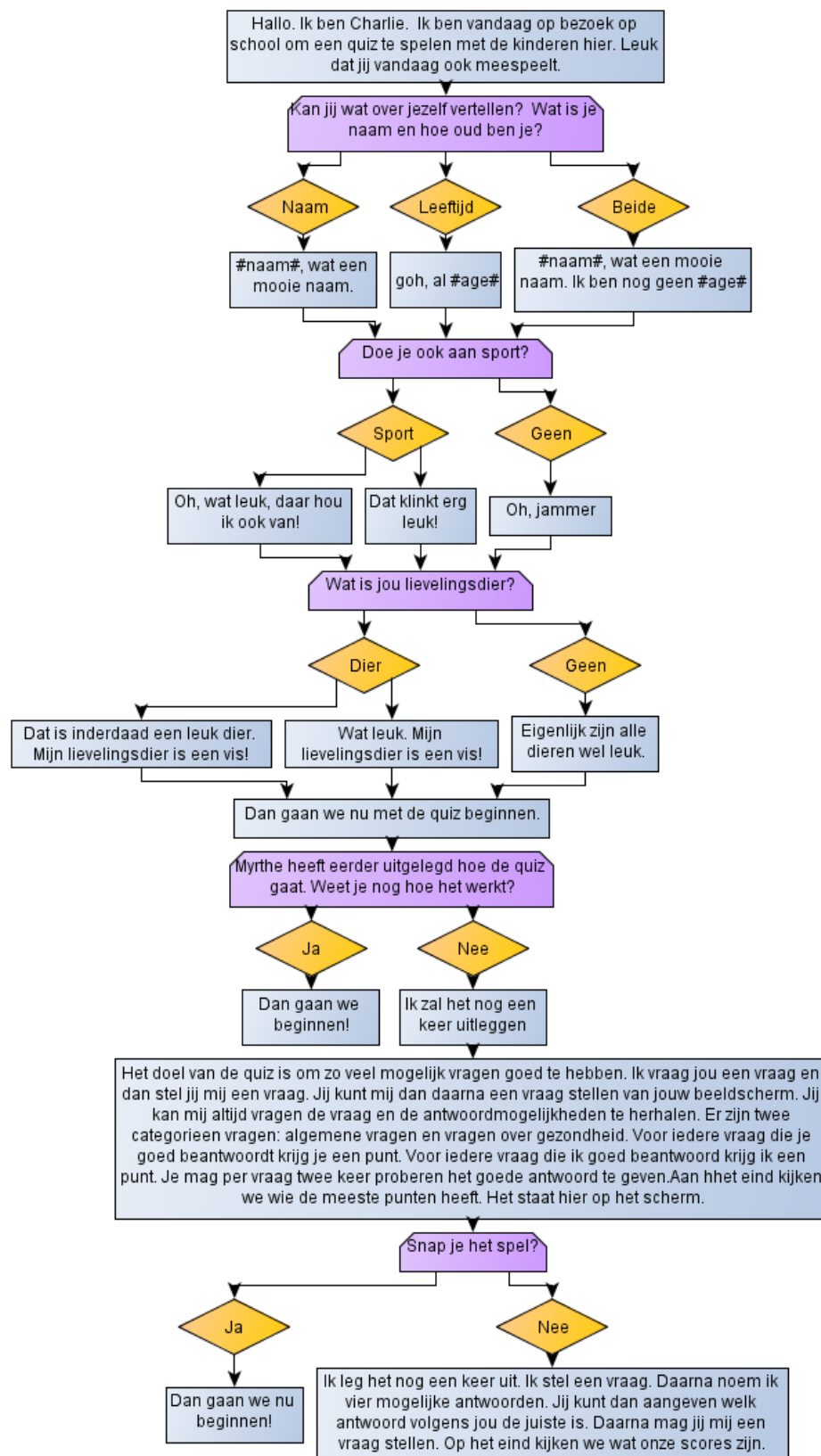


You will only need the tab which is open, the other ones can be ignored. You can adapt the valence and arousal by clicking somewhere in this space, each cross represents a combination of arousal and valence values. The X axis represents Valence and the Y axis represents Arousal.

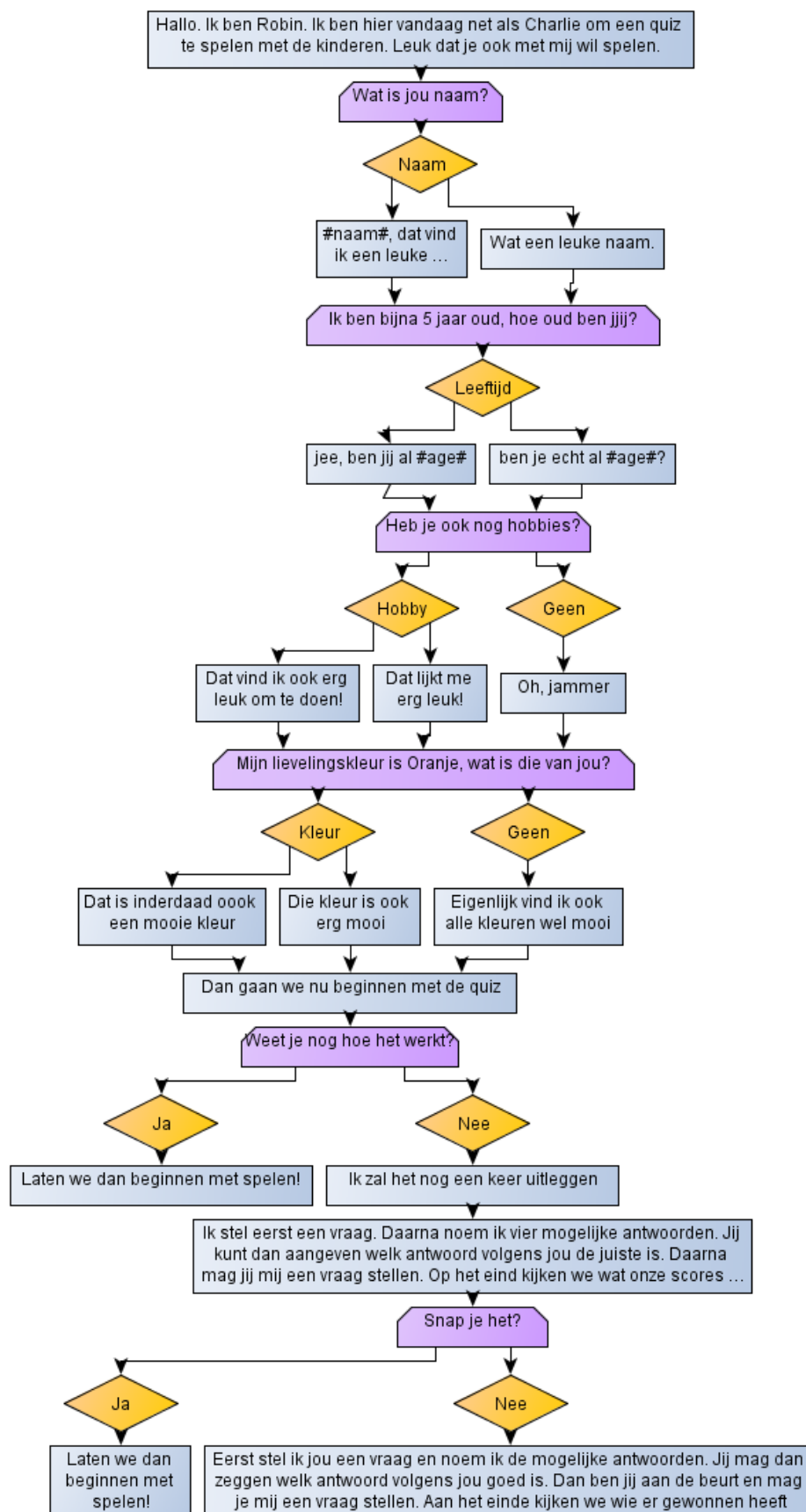
The black square is the place you have last clicked, so this should represent the current emotion of the child. (In this case it is Arousal 0.3, Valence 0.3; an emotion a little less positive than Glad). The coloured dots serve as reference points to specific emotions, so if the child seems to feel something between happiness and excitement, you should click somewhere between those dots. Every click will be rounded off to 0.1, so to one of the crosses. When the video starts, you may start interpreting the emotional state of the child. Try to do this as precisely as possible.

Appendix 7: Dialogs

A schematic overview of the dialogs used. Blue boxes represent remarks, purple boxes questions, yellow boxes possible answers from the child and green boxes quiz questions.



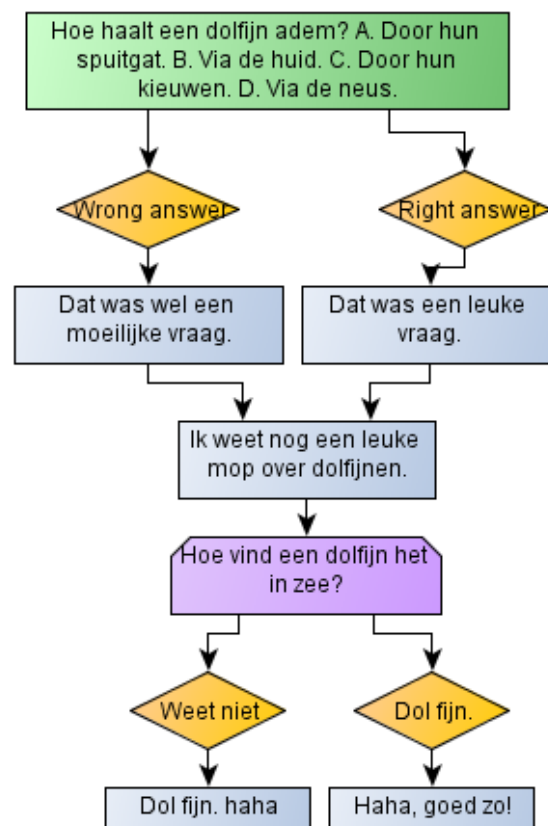
Introduction dialog first robot.



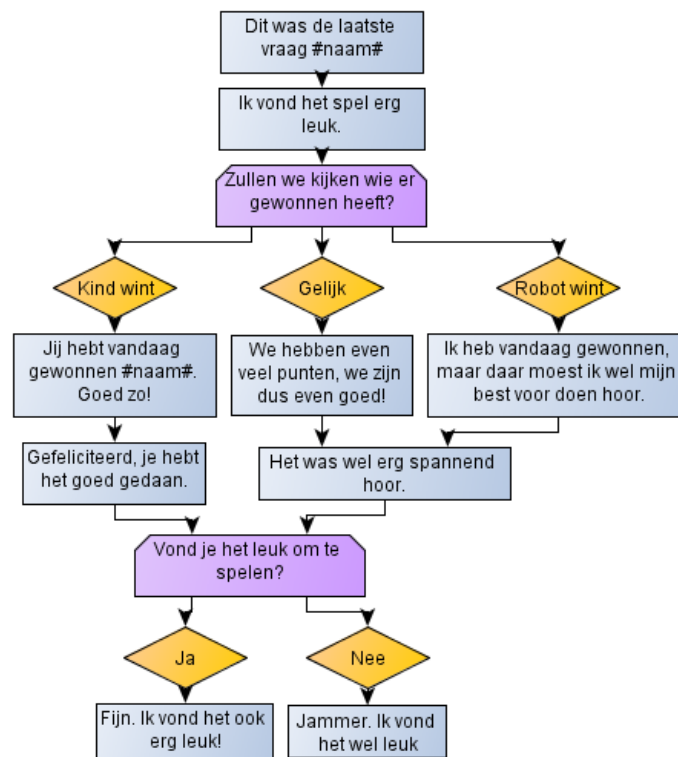
Introduction dialog second robot.



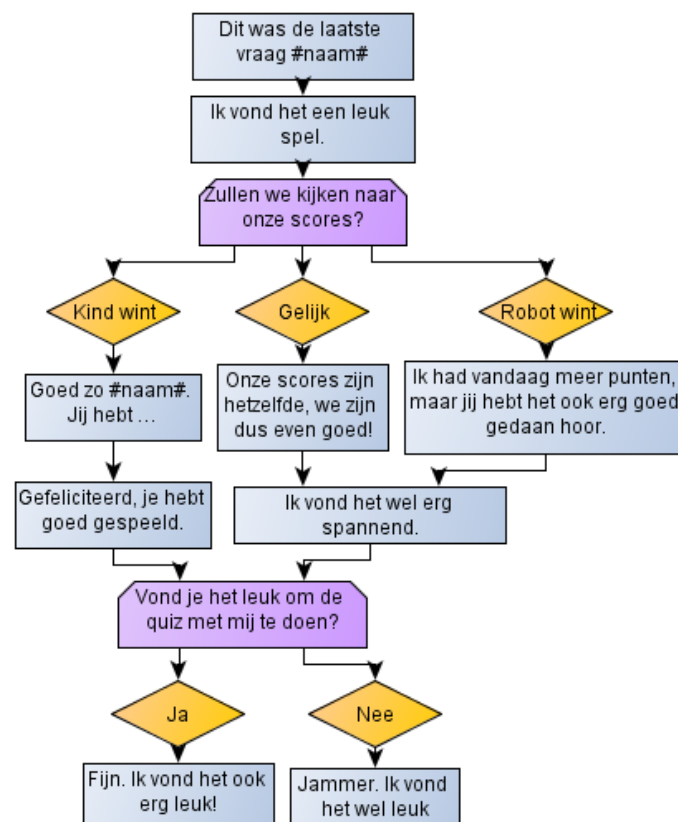
Joke half-way in the quiz for the first robot



Joke half-way in the quiz for the second robot



Concluding dialog first robot.



Concluding dialog second robot.

Appendix 8: Gestures

ID	Description
i1	Both hands go up and down again together. They stay in front of shoulders.
i2	Both hands go up and down again together. They are wider than shoulders.
i3	Left hand goes up, makes a fist and moves sharply down. (Success gesture)
i4	Right hand goes up, makes a fist and moves sharply down. (Success gesture)
i5	Both hands go up, make a fist and move sharply down. (Success gesture)
i6	Right hand goes up and down again, left hand goes up and down again. (One or the other gesture)
i7	Left hand goes up and down again, right hand goes up and down again. (One or the other gesture)
i8	Right hand goes up, left goes up, both hand go down. (One or the other gesture)
i9	Left hand goes up, right goes up, both hand go down. (One or the other gesture)

TABLE 8.1: Iconic gestures

ID	Description
b1	Left hand goes up and down, fingers closed.
b2	Left hand goes up and down, fingers open.
b3	Right hand goes up and down, fingers closed.
b4	Right hand goes up and down, fingers open.
b5	Left hand goes up, fingers closed.
b6	Left hand goes up, fingers open.
b7	Right hand goes up, fingers closed.
b8	Right hand goes up, fingers open.

TABLE 8.2: Beat gestures

ID	Description
d1	Right hand points to self.
d2	Left hand points to self.
d3	Right hand points to child.
d4	Left hand points to child.
d5	Right hand points to the lower right.
d6	Left hand points to the lower left.
d7	Right hand points to the upper right.
d8	Left hand points to the upper left.
d9	Left hand points to right.
d10	Right hand points to left.

TABLE 8.3: Deictic gestures

Appendix 9: Quiz Questions

The quiz questions asked in the order in which they appeared in the experiment. Question 1 through 12 were posed by the first robot, questions 13 through 24 by the second robot. The underlined answer is the correct one, note that this answer was not always presented first during the experiment.

1. Welke vrucht wordt als hij gedroogd is een rozijn?
<u>Een druif</u>
Een appel
Een peer
Een abrikoos

2. Hoeveel moet je elke dag bewegen?
<u>60 minuten</u>
15 minuten
30 minuten
120 minuten

3. Wat is het verschil tussen vlees en vis?
<u>In vis zitten goede vetten, in vlees minder</u>
In vlees zitten meer vitamines
Vlees kun je vervangen door vleesvervangers, vis niet
Vlees ziet er mooier uit op je bord

4. Hoeveel snaren heeft een gitaar normaal gesproken?
<u>Zes</u>
Drie
Twee
Vijf

5. Welke kleur heeft Spongebob?Geel

Blauw

Rood

Oranje

6. Wat zijn aderenDe vaten waar bloed door stroomt

Een slangensoort die in Friesland voorkomt

Stukjes van je maag

De kamers van het hart

7. Wat voor dier is een vleermuis?Zoogdier

Huisdier

Ongewerveld dier

Koudbloedig dier

8. Met welke sport is Epke Zonderland beroemd?Turnen

Hockey

Voetbal

Tennis

9. Kun je een vitaminepil nemen in plaats van groente en fruit?Nee, in een vitaminepil zit niet alles wat je nodig hebt

Ja, in een vitaminepil zit alles wat je nodig hebt

Nee, alleen volwassenen kunnen dat

Ja, als je een vitaminepil neemt die blauw is

10. Wat is gezond eten?Dat je genoeg van alle belangrijke voedselgroepen eet

Dat je niet veel eet

Dat je alleen maar groente eet

Dat je alleen maar snoep en snacks eet

11. Waarom is het belangrijk voor je hart om te bewegen?

Allemaal

Anders word je misschien dik en moet je hart harder werken

Anders zijn je hart en aders niet gewend om hard te werken

Anders worden je aders minder soepel

12. Afstanden op land geef je aan in kilometers. Wat gebruik je meestal op zee?

Mijlen

Inches

Voeten

Duimen

13. Waar zit je hart in je lichaam?

In je borstkas in het midden

Bij je ogen in de buurt

In je benen

Achter je navel

14. Welk land won in 2012 het junior Eurovisie songfestival?

Ukraine

België

Frankrijk

Italië

15. Hoe heet het volkslied van Nederland?

Het Wilhelmus

Oranje Boven

Holland Hoog

Nassau

16. Wat is gezonder, olie of boter?

Olie, dat zijn gezonde vetten

Boter, want dat is gemaakt van melk

Beide zijn even ongezonder

Beide zijn even gezond

17. Kun je ook te veel groenten en fruit eten?

Nee, je kunt er zoveel van eten als je wilt!

Ja, want sommige soorten zijn ongezond.

Ja, je mag maar 6 keer fruit per dag.

Ja, je mag maar 3 keer groente per dag

18. Van welke soort groente wordt Popeye supersterk?

Spinazie

Bonen

Aardappelen

Spruitjes

19. Is TV-kijken gezond?

Nee, maar ook niet ongezond, als je het maar niet te vaak doet

Nee, het is ongezond

Ja, het is heel gezond

Ja, als je maar goed op de bank gaat liggen

20. Wie is de keeper van het Nederlands elftal?

Stekelenburg

Edwin van der Sar

Hans van Breukelen

Piet Schrijvers

21. Hoe halen dolfijnen adem?

Door hun spuitgat

Via de huid

Door hun kieuwen

Via de neus

22. Waarvoor zorgen groenten en fruit?

Je helpt jezelf om er gezond oud mee te worden.

Je krijgt er direct een mooie huid van.

Je wordt nooit meer verkouden.

Je krijgt er heel veel energie van

23. Hoe heet de school waar je les krijgt in Dreuzelkunde?

<u>Zwijnstein</u>

Basisschool

Ramstein

Kostschool

24. Hoe belangrijk is het om te ontbijten?

<u>Ontbijt is elke dag heel belangrijk</u>
--

Je kunt het best overslaan, dan eet je gewoon meer bij de lunch

Ontbijten, daar word je alleen maar dik van.
--

Ontbijten, wat is dat?

Appendix 10: Questionnaires

Questionnaire introduction lesson

Je Leeftijd: ...

Ik ben een: Jongen / Meisje

1) Hoe leuk denk je dat het spelen met de robots wordt?



Vreselijk



Niet zo leuk



Leuk



Heel leuk



Geweldig

2) Heb je wel eens eerder met een robot gespeeld?

Ja

Nee

Questionnaire after each robot

This is the questionnaire for the first robot, the one for the second robot is identical, but with the name Robin in stead of Charlie.

1) Hoe leuk vond je het werken met Charlie?



Vreselijk



Niet zo leuk



Leuk



Heel leuk



Geweldig

2) Zou je Charlie nog een keer willen gebruiken?



Nee



Liever niet



Misschien



Ja



Heel graag

3) Zou je een keer een ander spel met Charlie willen spelen?



4) Zou je een keer een andere quiz met Charlie willen spelen?



5) Zou je een keer wat meer met Charlie willen praten?



6) Voel je je vrienden met Charlie?



7) Wat vond je het leukst aan Charlie?

8) Wat vond je het minst leuk aan Charlie?

9) Zou je Charlie thuis willen hebben?



10) Vond je het makkelijk om met Charlie te werken?



11) Begreep je Charlie?



12) Vind je Charlie aardig?



13) Denk je dat Charlie je begrijpt?



14) Vind je dat Charlie nieuwsgierig is naar jou?



15) Denk je dat Charlie de waarheid vertelt?



16) Zou je de vragen van Charlie eerlijk beantwoorden?



17) Denk je dat Charlie jou geheimen zou bewaren?



18) Denk je dat Charlie gevoelens heeft?



19) Kon je zien hoe Charlie zich voelde?



20) Denk je dat Charlie het leuk vond om met je te spelen?



Questionnaire comparing robots

1) Merkte je verschil tussen de robots?

- Ja
- Nee

2) Welke robot vond je leuker?

- Charlie
- Robin

Waardoor komt dat?

3) Welke robot vond je het makkelijkst te gebruiken?

- Charlie
- Robin

Waardoor komt dat?

4) Welke robot vond je aardiger?

- Charlie
- Robin

Waardoor komt dat?

5) Welke robot zou je het snelst een geheim aan vertellen?

Charlie

Robin

Waardoor komt dat?

6) Welke robot denk je dat de meeste gevoelens heeft?

Charlie

Robin

Waardoor komt dat?

7) Als je MOET kiezen, welke robot zou je dan kiezen?

Charlie

Robin

Waardoor komt dat?

Appendix 11: Correlations

The next page of this appendix shows a correlation matrix of different results from the experiment. From the questionnaire on each individual robot, the scores representing the differences between robots were used.

Correlations

	Age	Gender	Extroversion	Difference between fun scores	Difference between acceptance scores	Difference between empathy scores	Difference between trust scores	Difference between emotion scores	Preference final question	Score expressive robot final questionnaire	Score non-expressive robot final questionnaire	Difference between expressivity children	Difference between valence expressivity children
Age	1	-.412	-.350	-.006	-.555 [*]	-.156	.152	-.086	.106	-.301	.169	-.450	-.273
		.089	.155	.980	.021	.550	.562	.743	.707	.275	.546	.070	.288
	18	18	18	17	17	17	17	17	15	15	15	17	17
Gender	-.412	1	.317	-.079	.251	.254	.456	.496 [*]	-.055	.176	-.367	-.130	.261
	.089	.199	.199	.764	.332	.325	.066	.043	.847	.530	.179	.620	.312
	18	18	18	17	17	17	17	17	15	15	15	17	17
Extroversion	-.350	.317	1	-.027	.349	.174	-.026	.074	.211	.087	-.059	.273	.266
	.155	.199	.919	.169	.504	.504	.922	.778	.451	.758	.834	.288	.303
	18	18	18	17	17	17	17	17	15	15	15	17	17
Difference between fun scores	-.006	-.079	-.027	1	.506 [*]	.247	-.133	.460	.014	.215	-.312	.377	.482
	.980	.764	.919	.038	.611	.340	.611	.063	.960	.441	.258	.150	.059
	17	17	17	17	17	17	17	17	15	15	15	16	16
Difference between acceptance scores	-.555 [*]	.251	.349	.506 [*]	1	.472	.077	.491 [*]	-.066	.417	-.497	.251	.275
	.021	.332	.169	.038	.056	.770	.770	.045	.815	.122	.059	.349	.303
	17	17	17	17	17	17	17	17	15	15	15	16	16
Difference between empathy scores	-.156	.254	.174	.247	.472	1	.082	.184	.022	.255	-.146	.192	.311
	.550	.325	.504	.340	.056	.754	.480	.480	.939	.360	.605	.477	.242
	17	17	17	17	17	17	17	17	15	15	15	16	16
Difference between trust scores	.152	.456	-.026	-.133	.077	.082	1	.347	.524 [*]	-.346	.000	-.059	-.029
	.562	.066	.922	.611	.770	.754	.754	.173	.045	.206	1.000	.828	.915
	17	17	17	17	17	17	17	17	15	15	15	16	16
Difference between emotion scores	-.086	.496 [*]	.074	.460	.491 [*]	.184	.347	1	-.159	.431	-.500	-.102	.024
	.743	.043	.778	.063	.045	.480	.173	.173	.572	.108	.058	.706	.929
	17	17	17	17	17	17	17	17	15	15	15	16	16
Preference final question	.106	-.055	.211	.014	-.066	.022	.524 [*]	-.159	1	-.795 ^{**}	.641 [*]	.310	.227
	.707	.847	.451	.960	.815	.939	.045	.572	.000	.000	.010	.281	.435
	15	15	15	15	15	15	15	15	15	15	15	14	14
Score expressive robot final questionnaire	-.301	.176	.087	.215	.417	.255	-.346	.431	-.795 ^{**}	1	-.776 ^{**}	-.059	.006
	.275	.530	.758	.441	.122	.360	.206	.108	.000	.001	.001	.841	.985
	15	15	15	15	15	15	15	15	15	15	15	14	14
Score non-expressive robot final questionnaire	.169	-.367	-.059	-.312	-.497	-.146	.000	-.500	.641 [*]	-.776 ^{**}	1	.031	-.115
	.546	.179	.834	.258	.059	.605	1.000	.058	.010	.001	.001	.915	.695
	15	15	15	15	15	15	15	15	15	15	15	14	14
Difference between expressivity children	-.450	.130	.273	.377	.251	.192	-.059	-.102	.310	-.059	.031	1	.924 ^{**}
	.070	.620	.288	.150	.349	.477	.828	.706	.281	.841	.915	.17	.000
	17	17	17	16	16	16	16	16	14	14	14	17	17
Difference between valence expressivity children	-.273	.261	.266	.482	.275	.311	-.029	.024	.227	.006	-.115	.924 ^{**}	1
	.288	.312	.303	.059	.303	.242	.915	.929	.435	.985	.695	.000	.000
	17	17	17	16	16	16	16	16	14	14	14	17	17

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).