AALBORG UNIVERSITY

Social Robots: Comparing Young and Old Participant's Perception of Affective Body Movement

by

Anders Krogsager & Nicolaj Segato

A thesis submitted in partial fulfillment for the degree of Master of Science

in the
Faculty of Engineering and Science
School of Information and Communication Technology

Wednesday 4th June, 2014

"Her i Troldspejlet interessere vi os jo mest for de sjovere ting ved computere. Sådan noget som regnskaber, statistikker og beregninger og den slags skal man ærligt talt være en lille smule mærkelig for at kunne se noget som helst spænedende i."

Jakob Stegelmann
Troldspejlet, 2. November 1989

Acknowledgements

Special thanks go to the following individuals: Professor Matthias Rehm for countless hours of supervision, constructive feedback and several thick books. To the employees and volunteers at institutions around Aalborg, there would have been no tests without their help: Bent Sørensen for welcoming and accommodating the project at Fremtidens Plejehjem. Leif Fuglsang for his tireless efforts in recruiting for and supporting tests at Fremtidens Plejehjem. Lone Svendsen and the other employees at Liselund Aktivitets-center for their efforts in recruiting participants and for allowing us to conduct tests in Vodskov. Sara Hedegaard, Anette Juhl, and the other employees at SOSU Nord for allowing us several visits to tests with students. To Rasmus Stenholt, for sharing his statistics knowledge.

Contents

List of Figures

List of Tables

Abbreviations

ABM	**A**ffective **B**ody **M**ovement
HRI	**H**uman **R**obot **I**nteraction
RQ	**R**esearch **Q**uestion
SD	**S**tandard **D**eviation
M	**M**ean
Mdn	**M**edian
KM	**K-M**eans
VD	**V**ascular **D**ementia
EEG	**E**lectro**E**ncephalo**G**raphy
fMRI	**f**unctional **M**agnetic **R**esonancec **I**maging

Chapter 1

Introduction

As robots starts taking on more ubiquitous roles in society, they must be easy for everyone to use and interact with, whatever the users are young or old. They have to interact with the user in a socially acceptable manner, so people feel comfortable interacting with these robots. Humans have a natural tendency to interpret upon stimuli they perceive, once something is perceived it is categorised and organized, then it is superimposed upon our own personality. Because humans perceive according to where they are in their lives, gives rise to an important question of how young and old perceive emotions expressed through a robot? How are these emotions perceived in laboratory conditions vs. a context?

1.1 The Robot Interface

Social humanoid robots pose a dramatic shift in the way people think about how the workforce of the future is going to operate. Traditionally, robots have been designed to be operated directly or as independently and remotely form humans as possible can be. Often robots perform hazardous task in hostile environments e.g. sweeping mind fields, decontaminating nuclear waste or exploring other planets. Applications like mowing lawns, hoovering or functioning as courier at hospitals is still minimal. However various industries drive the development of a new range of robot application domains i.e. domestic, health care, surveillance etc. that can cooperate and interact with people as a collage rather than functioning as a tool. Within the field of human computer interaction (HCI) research Reeves et al.[22] showed that humans (whether computer novices or computer experts) often treat computers as they might treat other humans. Furthermore from their study [22] argue that a social interface could be thought of as a universal interface, which arguably a humanoid robot or virtual agent is. Sharing a comparable morphology with humans, humanoid robots can communicate in a way that supports natural human communication. e.g. using gestures or express emotions through facial expressions. As such a humanoid robot can interact and work along with users in a way that supports this natural way of communicating and can with more ease be integrated along the workpool within e.g. health care. The increasingly skewed distribution between young and elderly in technology societies such as Denmark will need a larger workforce to care for those who are weakened by age. Social robot technology is expected to alleviate part of this task.

1.2 Motivation

Previous research by Krogsager, Segato and Rehm [1] describes Danes interacting with a humanoid robot in a laboratory setting. Users are asked to speak of a topic to a robot in a first encounter. It is then measured how the head movement of the robot influences the duration of speech. The participants were expected to behave towards the robot in accordance to social rules but test results shows this is not the case. While the robot reacted according to social rules users did not respond as expected. A possible explanation is the lack of a social context for the interaction in a laboratory setting. If participants do not understand the situation they cannot react appropriately.

One possibility is to set up a game scenario to create context for emotional responses. The advantage of a game is that an emotional response happens towards the game and not a partner of interaction. A test should be carried out to investigate how Danes react to a robot in the context of active interaction such as a game. Seniors are likely to be the early adapters of commercial social robots, but research by Nomura et al. [20] shows that basic emotion recognition in Japanese seniors differ from that of the young generation. While interaction with robots is more likely developed and tested by young generations, interaction can still break down due to age difference. Before assuming that Danish users will interpret emotions differently due to age this must be tested since the age-dependent emotion perception may be reliant culturally on Japanese. The experiment shows how affective body movement in a humanoid robot is not perceived alike by students and seniors. If Danish seniors bear the same perceptual impairment then robot interaction design should take into consideration how to avoid ambigous behavior in humanoid robots.

1.3 Research Questions

Nomura et al. [20] poses two reqearch questions: "1. Are there differences on emotion identification of body motions expressed by robots between younger and elder people?" "2. Are there correlations between the accuracy of this emotion identification and cognitive bias to robots specific body motion parts?"
which are combined into RQ1 in this project. The first objective is to replicate the study with Danish participants. RQ2 is asked to clarify if the presence of a robot in the experiment [1] is the cause of unexpected results compared to its source experiment. Research questions 1 and 2 regards the ability of seniors to interpret emotions from a robot. RQ3 is investigated with a game between a robot and a senior and should explain if context to an expression of emotion changes perception of it.

1. Do Danish participants interpret the movement behavior similar to Japanese?
2. Do Danish participants perceive affective body movement more accurately with a co-located robot?
3. Does contextualization of affective body movement improve accuracy of perceived emotions?

Chapter 2

Background

In this chapter a range of implicated topics are presented. Each section describes related research and state of the art in the fields of relevance to answer RQ1, 2 and 3.

2.1 Emotions Overview

There are two main schools within emotions [10] one that promotes a categorical approach, which define several distinct emotions such as anger, happiness, fear, sadness and disgust. The other school has an dimensional approach in Figure 2.1. Approaches from theorists in various fields including psychology, phylogenetic and psychophysiology have proposed different frameworks influenced by the insight of their respective fields. The lack of a clear definition of emotions can be attributed to the subjective and untangible nature of emotions [13]. These studies relied upon ”semi-ethnographic” methods [9] or are confined to either using ”noisy” data such as self-reported metrics. Other studies on emotions relies upon equipment like electroencephalography (EEG), precious scanning methods such as functional magnetic resonance imaging (fMRI) or invasive methods similar to as single-unit recording [10].

2.1.1 Nominal and Ratio Approaches to Emotions

In [21] (chapter 3) Paul Ekman cite Stein and Tarbasso (1992) for their description of the emotions (anger, happiness, sadness) which is applied in many related HRI studies including [32] and [20]. The quote well exemplifies the function of emotions and is also brought here:

> ...in happiness a goal is attained, in sadness there is a failure to attain or maintain a goal, in anger an agent causes a loss of a goal...

The theory of nominal classification promoted by Paul Ekman is frequently applied in human robot interaction research such as [3, 5, 14, 20, 29].

The theory of psychophysiology initiated by Charles Darwin[26] [15] states that emotions have evolved to increase survival chances by affecting behavior as they trigger physiological responses. Emotions also serve as signals for future actions and intentions to individuals in the surroundings. This is further supported by Ekman [18], [9] and [21] which states that emotions are a biological product of evolution and functions as an integral part of the survival mechanism. Therefore emotions must be universal and anger should be experienced in the same way in Japan as in Denmark or any other place in the world. This is supported by research conducted by Ekman [9] in the late 1950's to the early 1960's where Carltob Gajdusek [26] provided film of two preliterate New Guinea cultures. No unique facial expressions were found and this is used as an argument to the question of whether facial expressions are socially learnt or innate, evolved behavior. Additionally, Ekman suggests that the primary function of emotions is to enable one to deal with interpersonal encounters based on past experiences. The notion of basic emotions implies that complex emotions can be formed by combining basic emotions. In [21] Ekman lists the following fifteen as basic emotions: amusement, anger, contempt, contentment, disgust, embarrassment, excitement, fear, guilt, pride in achievement, relief, sadness/distress, satisfaction, sensory pleasure and shame. These 15 are adhere to a list of eleven characteristics that define them, to mention a some: distinctive universal signals, distinct physiology (physiological reaction), presence in other primates, brief duration, unbidden occurrence, etc. The fifteen emotions are debated and the requirements for what constitutes an emotions varies between researchers. Ekman's basic emotions are meant to highlight a gap in knowledge about emotions and are expected to change. In opposition to the categorical approach favored by Ekman and others, some researchers argue for a multi-dimensional framework. The two-dimensional variety are positive/negative affect and high/low arousal, which some find to be a more appropriate definition of emotions See Figure 2.1. The advantage is that all emotions be defined from these two dimensions rather than restiricting them to a categorical definition. Emotions then also overlap on the two dimensions. As an example anger can be defined as high arousal and negative affect, but so can fear [10].

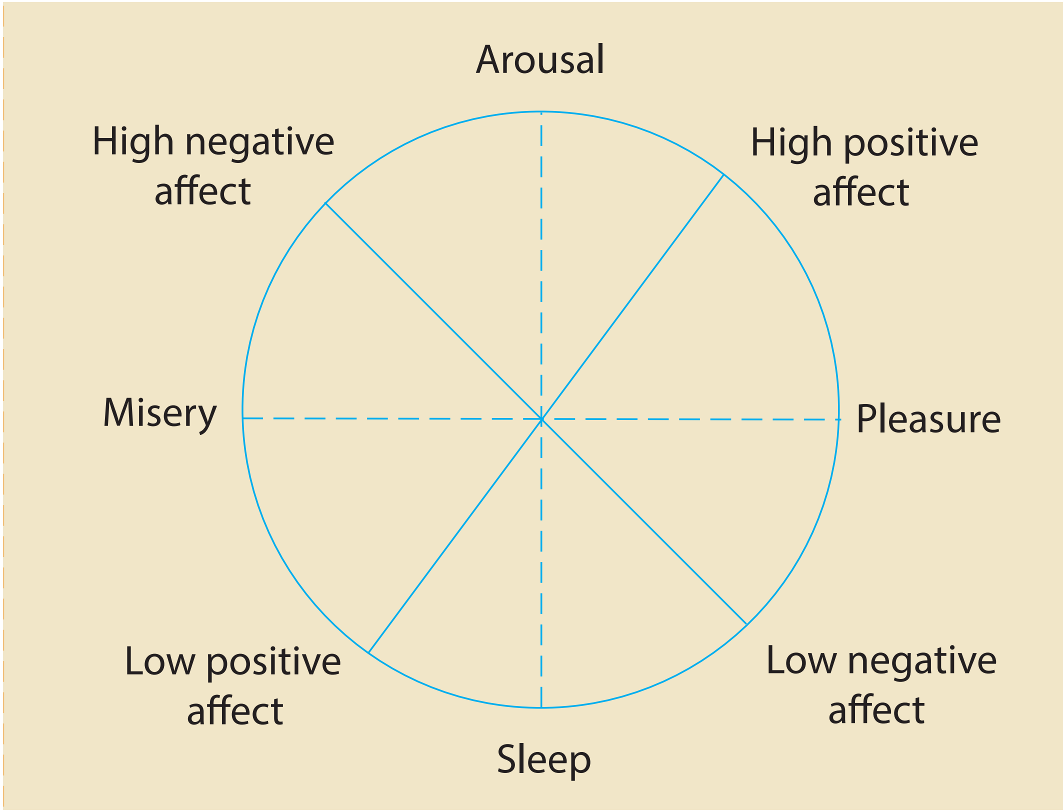

FIGURE 2.1: [10] Shows a two-dimensional framework for emotions. The two dimensions of positive affect and negative affect and the two dimensions of pleasure-misery and arousal-sleep. Emotions like excited and happy fall within the top-right quadrant, while e.g. calm, relax and contented fall within the bottom-right quadrant, and so on.

As support to the ratio approach Shaver et al.[15] used a quantitative method to determine a set of basic emotions using participants who had to compare 135 emotion terms; these emotion terms where based on empirical self-reported data generated from experiments on emotion terms that was founded by research of emotions. The study used hierarchical cluster analysis to analyse the gathered data, and categorise certain emotions into certain groups. The study states that people identify the same emotion terms; six major terms were determined, joy, love, sadness, fear, surprise. [15] never mentions the last emotion. This project will not discuss emotion theories but simply apply them and acknowledge that there are many different theories and that the ones we utilize is not considered *ne plus ultra*.

2.1.2 Influence of Emotion

Plutchik [21] Define the term emotions "as a patterned bodily reaction of either destruction, reproduction, incorporation, orientation, protection, deprivation, rejection, or exploration, or some of these, which is brought about by a stimulus". Emotions works as a automatic apprising system that is continually scanning our world, it detects when something of importance to us is happening e.g. our welfare or our survival. Emotions do not only help us in survival situations, they also change how we see the world and interpret the actions of others. Thus emotions are a governing aspect which can determine the quality of our lives as they occur all the time throughout our whole life [9]. Emotions are expressed e.g. involuntarily through facial expression, tone of voice and also posture and gesticulation. However depending on social situation and cultural context the expression of a given emotion may be frowned upon by others, such as negative affect in public in Japan. As such the emotion is repressed. Since emotions may be masked or faked a layer of complexity is added [26]. The research on emotions shows that emotions are intuitively produced and recognized and also allows the assumption that basic emotions are understood across cultures.

2.1.3 Dementia, Age and Emotion Perception

Sejerøe-Szatkowski [8] states that through body language humans produce information about our state of mind. This information differs from verbal information, in that it is often emitted unconsciously by people. To people suffering from dementia symptoms, body language is often relied upon more than it is in healthy individuals, and as such can lead to misunderstanding. This increased dependency is a result of breaking temporal continuity due too reduced short-term memory. Therefore body language and facial mimic bears great important as a channel of expression and information reception to afflicted individuals. The decline in cognition, emotion recognition and expression is caused by the degeneration of important signal transmitters in the brain [4].

Research by Wong et al. [32] showed that emotional perception declines not only in patients that suffer from dementia, but in fact do with the increase of age. With age a difference in ocular scanning patterns (eye movement patterns) of human faces occurs which affects emotion perception. In their study participants were shown images of faces and listened to recordings of voices. Older adults had a lower accuracy in identifying emotions of anger, fear and sadness in the images compared to young adults. These results are consistent with their sources that compares the emotion identification of old and young participants. The affective signals were generated by showing black-and-white photographs and playing audio clips of speech. The group of older adult participants had greater difficulties in accurately identifying anger, sadness and fear in pictures. It should be noted here that the results were different for the accuracy of emotion identification in voices. This shows that emotion perception in one channel does not mean unanimously impaired perception. The decline is attributed to neural loss and habituation to emotions through life experiences. The two groups were compared in an anti-saccade task [1]where it was hypothesized that emotion identification would correlate with task performance, yet the results were insignificant. In the final test the duration of fixation on a photo was measured using an eye-tracker. The results showed that the older adults made fewer fixations and did so lower down on faces, which impaired some emotion identification.

2.1.4 Application of Emotions Theory

Here, as in the experiment of Nomura [20] emotions are decontextualized. This means there is no reference to a facial expression to accompany the movement or vocalizations to indicate affective state of the robot. Nor is there a scenario in a social context as an indication to what emotion might be expressed. This is an unnatural condition which makes it more difficult for the person observing the robot to identify an emotion. This disadvantage of decontextualized emotions is accommodated for by accepting a degree of inaccuracy of responses in upcoming tests. Following the two-dimensional model of emotions affective signals are represented as points on high/low arousal, positive/negative scales. Shifting from nominal to an ordinal affect representation (e.g. anger to high arousal, negative affect) an emotion may be recognized correctly. If an emotion is identified with a different affective term, e.g. "hate", when the emotion of anger was displayed then this can be considered a correct identification.

[1]In an anti saccade task, a visual cue is presented to either the right or left of the participant. The participants role involve not looking at the presented cue but, rather inhibiting the response and look in the opposite direction.

2.2 Human-Robot Interaction and Emotion Research

The main work that lays foundation for this project is described in section 2.2.1 and is also referred to as the Nomura experiment. Subsequent sections describe other related works in human-robot interaction, from which methods are applied in this project, aside from gesture annotation.

2.2.1 The Nomura et al. Experiment

Nomura et al. [20] is a study on the influence of age in perceiving emotions from a humanoid robot. The study uses two different age groups (young and senior participants) and investigates how the participants perceives the emotions (anger, sadness and pleasure) expressed by the robot. The results of Nomura are reached by conducting a laboratory experiment where whole body movement is produced by a humanoid robot and participants are asked to report their impressions of each movement. Experiments by Wong et al. [32], and their sources, on age differences suggests that seniors have lower accuracy in identifying anger, fear and sadness compared to young participants. In [20] the emotion "pleasure" is chosen as a positive identifiable contrast to the negative emotions, anger and sadness. In addition Nomura recognize that the accuracy of the emotion identification in seniors could be affected by cognitive bias to specific areas of the robot, which is very similar to the explanation made by Wong. While Wong et al. test emotion perception from voice and images of faces, Nomura tests whole body movement. Regardless of this change in procedure Nomura also finds that there is a difference in emotion perception between young and seniors. Nomura looks into which regions of the robot, specifically which body parts, people look at; e.g. arms, legs, head etc. during the session. Though this method may be an inaccurate method and be considered questionable. Attention is partially an autonomous process [28] which makes self-reported measures of attention difficult.

2.2.2 Human-Human Interaction

The basis for the Nomura experiment is [32], but Wong et al. do not focus on body movement, but rather faces. A prerequisite for this project is the ability in humans to perceive emotions in other humans through body movement alone. Several studies such as [6] and [31] have investigated how people perceive affective body movement, "ABM", (and not gestures) in others, though "movement" is not just "movement". In conversation, gestures (iconic, symbolic, metaphoric, etc.) can give redundant(extra/repeated) and additional information. The work of McNeill [17] contribute a proficient framework and annotation scheme for identifying the gestures of a conversational partner. The gestures McNeill describe occur intentionally to emphasize a part of a sentense or visualize it. These gestures allow the speaker to visualize what he has experience (seen or heard) in the context of his verbal explaination. These gestures are learned and culturally dependent, in other words acquired though life experiences and not an evolutionary matter of course. Gestures outside conversation, as seen in [20], are of another nature. Various dimensions for annotating the movement are described between researchers such as Walbott and Meijer [2] and one, Laban [16], is frequently used in the HRI research, including sources [23], [19], [20].

Gestures are often researched and described in the context of conversation, such as McNeill [17], wherein gestures are used to represent the shape and movement of physical things or metaphores. For example a person describes a stone being thrown through the air. The speaker then uses his hand to symbolize the movement of the stone flying in an arch and the speed or force of it. The emotion display in ABM is isolated from deliberate purpose in social interaction and represents a mental state rather than the motion of a physical thing.

These emotions have previously been looked into in the existing studies on emotion expression by robots. Nomura et al. and Wong et al. used six basic emotions; fear, pleasure, anger, surprise, sorrow and hate (Nomura) / disgust (Wong), in addition Wong et al. used a natural emotion was used in this study. Because of cultural difference the emotional perception of movement differs and there is no guarantee that the meaning will translates from Japanese to Danish culture [23]. In order to answer the research questions in this project it is not possible to simply re-use the movement used in [20]. In many experiments, including [32], [6] and [31], an actor is the source of ABM but results are then biased by the performance of the actor and his/her subjective expression. Their ABM does not reflect a felt emotion but rather a "simulation" of the emotion. Generating movement for a single "unmixed" emotion is not be possible as emotions are not felt unadulterated [13]. Though it will not be pursued in this project an alternative to using actors could be to review recordings of multiple individuals in situations of real emotional experience. Gross et al. [12] attempted to elicit emotions in actors and then record movement for the task of knocking a door but also conclude that emotions do not come unadulterated. Rehm [24] takes the approach of user-centered design to generate ABM. This approach has a clear advantage over actor interpretation in that it relies on a group impression of what emotion is expressed on given parameters of movement. [24] defines movement by the annotation scheme of Rudolf Laban [16] whose work is prevalent in HRI research. It also makes movement ranslateable to other robots.

Research by Meijer [6] suggest that ABM can be understood across culture based on evolutionary psychology and its influence on movement. For example sadness and anger may be conveyed through gait [30]. Researchers have already found certain features in movement that relate to specific emotions [2] which are compatible and to some extent overlap with the basic emotions of Ekman and the multidimensional representation of emotions (2 or 3 axis). Working with the assumption that humans unconsciously perceive emotions in all movement, as it seems to be the case in evolutionary psychology, then the paper by Goetz [11] finds an application by taking personality into consideration in robots. They showed how the adapted personality made cooperation between human and robot more effective. Cynthia Breazeal [3] applied the multidimensional representation of emotions to simulate affective states in the social robot Kismet, demonstrating the practical application for this theory in human robotics.

2.2.3 Robot Appearance

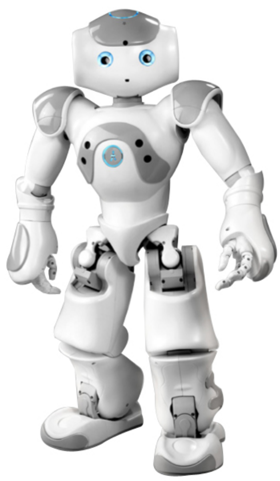

Figure 2.2: Robot appearance used in the project.

It is well established that embodiment of an agent or robot affects perception. The robot in this project is with relation to studies in anthropomorphism and the uncanny valley not at risk of causing displeasure from having close physical resemblance to humans. It is assumed that if participants feel uncomfortable while testing with the robot their perception feedback will be affected, thus skewing results. See figure 2.2

Chapter 3

Affective Body Movement Design

The design, implementation and verification of affective body movement for the robot is described in this chapter. The ABM is then used in subsequent experiment chapters.

Previous studies in human-human interaction and HRI as described in section 2.2.2 employ actors to produce body movement. The disadvantage of this approach is then that movement is based on the subjective expectations of that given actor to express distinct emotions, so the movement is not necessarily generalizable for a large demographic. ABM in the Nomura experiment [20] is inspired by Japanese Bunkai puppeteering. There is no equivalent Danish puppeteering tradition so for ABM in this project an alternative, nationally recognizable source of ABM is selected: The design of ABM for the robot is inspired from a Danish TV show titled "Bamses Billedbog" produced for children. The advantage of replicating the actors of this show is their exaggerated body movement as they wear full-body costumes with little or no facial animation. The Nao robot has no features of facial animation and is not equipped to perform minute movements. Therefore the performance of the given actors is a suitable match for the robot.

Three clips containing ABMs are selected based on several hours of review of the show. They are chosen for their distinct expression of anger, sadness and happiness and the clips are located on the enclosed DVD in the folder 'ABM design\Actor ABM source videos'. Figure 3.1 shows series of screenshots from the three clips. This ABM selection is subjective so to ensure that the emotion expressions of anger, sadness and happiness are not dependent on their context in the program a survey is conducted which is described in section 3.2.

FIGURE 3.1: Screenshots from the three videos with in individual displays of sadness, happiness and anger. First row shows the clip of sadness, second row shows happiness, third row shows anger.

3.1 ABM Implemention

This section describes the practical process of implementing ABM in the Nao robot. The process is fairly intuitive and uses the software suite Choregraphe bundled with the robot. Animation of the robot is handled by placing the robot's body in keyframe positions from where the software interpolates between the positions. The process was done by looping the sequence of the actor and matching key positions by posing the robot.

It is important to notice that noise is created from this translation process between an actor on video to a physical robot. An additional survey is conducted to assess how the translation affects the perceived emotion by viewers. Figure 3.2 shows the three ABMs implemented in the robot. Videos of the movement can be found on the enclosed DVD in the 'ABM design' folder.

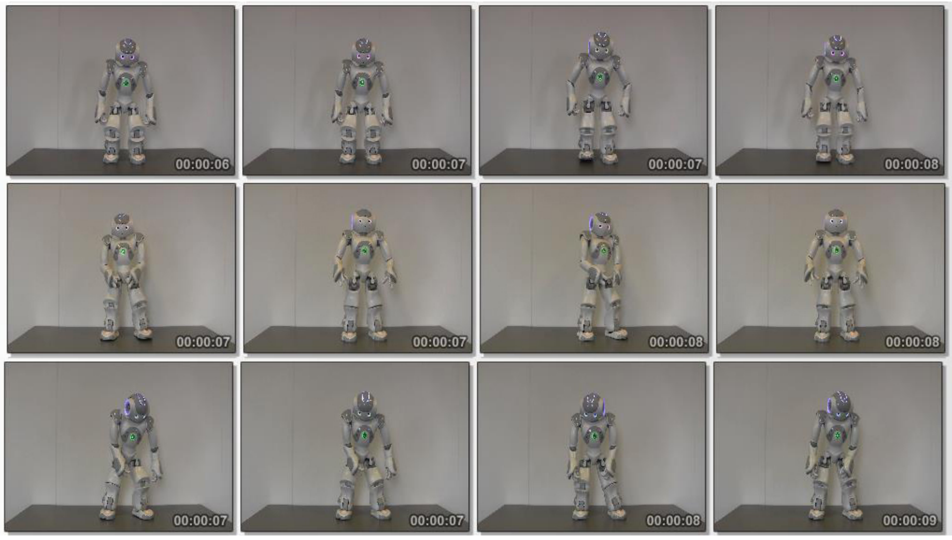

FIGURE 3.2: Images of the three animations with in individual displays of anger, happiness and sadness. First row shows the clip of anger, second row shows happiness, third row shows sadness.

3.2 ABM verification Survey

A preliminary test is conducted to ensure validity of the selected ABM from the TV program. In context of the program the emotional signal of the actors movement is enhanced by narrative context and sound. The test is to ensure that the three body movements are perceived to convey the same emotions when decontextualized from the TV program. If the body motion of the actor is interpreted as another emotion than the one intended, then other motions must be selected for the robot to express. Participants are shown the three selected clips in random order and asked what emotions they think the character expressed. Each clip has a duration of 5 seconds, has no sound and is shown without context. After an extended period of time (approximately 3 months) the participants are surveyed again with video of the robot expressing the same three ABMs. The design of the survey generally follows the same design as the experiment of emotion identification in [20]. Participants must be able to identify the correct emotion with a higher accuracy than by chance. The emotions on the list follow six of Ekman's basic emotions; anger, hate, joy ("happiness" used interchangeably), surprise, sadness and fear. An additional option of 'other' is present. Each item has three answers: one strong positive, one weak positive and one strong negative. The reason for this asymmetrical distribution is to prevent a false negative response, and is in line with the procedure of the source experiment in [20]. Participants are given the feedback options of "I am certain", "I think so" and "I do not think so" in regards to what emotions the character in the video seemed to express. Since participants are allowed to report perceiving several emotion it is possible that they report emotions of both positive and negative affect in the same video. It does not void the validity of a correctly identified emotion if the participant also reports perceiving another emotion of opposite affect, since it is assumed that perception is enhanced once ABM is contextualized. Fisher's p is calculated to show if a statistically significant difference between robot and actor is present. Ideally the robot performs equally or better than the actor in conveying the three emotions.

3.2.1 Participants

Participants are Danish students who are expected to be familiar with the TV program due to its long run time and popularity. No special requirements are set for participants to be accepted into the survey. After the survey participants are asked if they are familiar with the program, their age and their field of study (recorded at first session of the survey). The survey is conducted in the participants home, a dormitory in Aalborg, Denmark.

3.2.2 Equipment

A laptop with a 17" screen is used to display videos and collect data. Sound is disabled.

3.2.3 Procedure

Each participants is asked to see the three video clips and after each report the emotions they recognized from a list. The setup of this list is almost identical to the first page on in Appendix A. Participants were shown videos in random order. They were allowed to replay each video multiple times. The list of emotions is modeled after the one by Nomura [20] and participants can select several emotions from it. If a participant choose "other" they are asked to describe another emotion they thought the character expressed. Sound was disabled during the test.

3.3 Results

Participants (N = 11) were students or recently graduated of Danish nationality, aged 20-26 (M = 23,2. Std = 2,6. Mdn = 24), 7 males and 4 females. Their field of study was varied and include psychology, communication, politics, economy, law and engineering. All participants except for one were familiar with the TV show from when they were young.

The results of the 11 participants are show in the tables 3.1 3.2 and 3.3 compares the identified emotions from actor and robot. Each table represents a video of one distinct ABM and participants could choose multiple emotions. The percentage indicates how many of the participants expressed perceiving a given emotion. p value reported for Fisher's exact test is 2-sided.

Results for ABM1 (anger)			
	Robot	Actor	Fisher's p
Anger	11 (100%)	11 (100%)	*
Hate	8 (72,7%)	3 (27,3)%	0.086
Happy	0	0	
Surprise	0	0	
Sad	1 (9,1%)	1 (9,1%)	1,000
Fear	1 (9,1%)	0	1,000
Other	3[1] (27,3%)	3[2] (27,3%)	1,000

Table 3.1: Summary of questionnaire results from the test video showing anger. It shows the number of participants who perceived a given emotion for each video. *Chi-Square not calculated.

Results for ABM2 (happiness)			
	Robot	Actor	Fisher's p
Anger	0	3 (27,3%)	0,214
Hate	0	0	
Happy	10 (90,9%)	9 (81,8%)	1.000
Surprise	5 (45,5%)	1 (9,1%)	0,149
Sad	0	1 (9,1%)	1,000
Fear	0	0	
Other	1 (9,1%)	2 (18,2%)	1,000

Table 3.2: Summary of questionnaire results from the test video showing happiness. It shows the number of participants who perceived a given emotion for each video.

[1] regret, frustration
[2] frustration, irritation, impatience

Results for ABM3 (sadness)			
	Robot	Actor	Fisher's p
Anger	0	3 (27,3%)	0,214
Hate	0	0	
Happy	1 (9,1%)	0	1,000
Surprise	1 (9,1%)	0	1,000
Sad	9 (81,8%)	6 (54,5%)	0,361
Fear	3 (27,3%)	0	0,214
Other	4[3] (36,4%)	8[4] (72,7%)	0,198

Table 3.3: Summary of questionnaire results from the test video showing sadness. It shows the number of participants who perceived a given emotion for each video.

3.4 Analysis

The translation of affective body movement from actor to robot is considered successful. For the three ABM expressions there is no significant difference in perceived emotion from actor to robot. For ABM1 (anger) in table 3.1 all 11 participants perceived anger. For ABM2 (happiness) in table 3.2 10 participants perceived happiness in the robot which is higher than the rate for the actor. For ABM3 (sadness) in table 3.3 9 participants perceived sadness in the robot while only 6 reported seeing this emotion in the actor. In conclusion the three ABM animations in the robot are considered usable for further testing.

[3]embarrassed

[4]disappointment, frustration, regret,irritation, despair, confusion

Chapter 4

Experiment 1

This chapter summarizes the design, execution and results of experiment 1 which is a replication of the 2010 study in [20] described in section 2.2.1, now under a different cultural condition. This experiment aims to answer research questions 1 and 2. Before pursuing RQ3, it must be established whether Danish seniors have the same reduced accuracy in emotion recognition in humanoid robots as seen in Nomura's experiment. In summary; it is investigated if senior Danes have lower emotion recognition accuracy than young Danes, as the case is for Japanese [20]. In that case general design principles for affective signalling in HRI cannot be applied ubiquitous to young and old users. That calls for enhanced ABM design requirements. Experiments are conducted on-location and steps are taken to keep the room clear of exogenous cues and for distractions. To the extent possible the wall in front of the participant is cleared and a do-not-disturb sign is placed on the door. All experiment recordings and collected documents are found on enclosed DVDs labelled 'Experimental documentation'.

4.1 Experiment Design

The purpose of experiment 1 is to replicate Nomuras experiment before investigating RQ3. This repeated-measures experiment presents participants with three different affective body movements from a humanoid robot. The three expressed ABMs are, as in Nomuras experiment, anger, sadness and pleasure. The participants are co-located with the robot under 'laboratory' conditions. The hypotheses are derived from the research questions listed in chapter 1:

Hypothesis 1*. Danish seniors will have lower accuracy of emotion perception than young.*

Hypothesis 2*. The presence of the robot affects emotion recognition accuracy of both groups of participants.*

Experiment 1 has two groups (young - seniors) and two conditions (co-located - virtual presence) which gives four combinations of results. The table 4.1 shows a summary of variables in experiment 1.

Independent variables

Variable	Method
AMB affective signal	Distinct movement of the robots body; anger, sadness and pleasure
Presence	Displaying the robot physically or on a screen

Dependent variables

Variable	Measure
Affective signal perception	Participant feedback of the emotion they recognized
Magnitude/speed perception	Participant feedback of the movement they recognized
Interview question	Participant opinion/idea of use of a helper robot
Fixation location	Eye tracking of area where participants look

Table 4.1: Summary of variables in experiment 1. The independent variables are controlled by the test conductors. The dependent variables are the measured outcome of the experiment.

The measures in the experiment are based upon a self-reported method and data gathered by the eye-tracker.

The questionnaire consisted of two parts, the first part is a group of items designed to measure which emotion the participant perceived in the expressed ABM. As in the test in section 3.2 each item has three answers: one strong positive, one weak positive

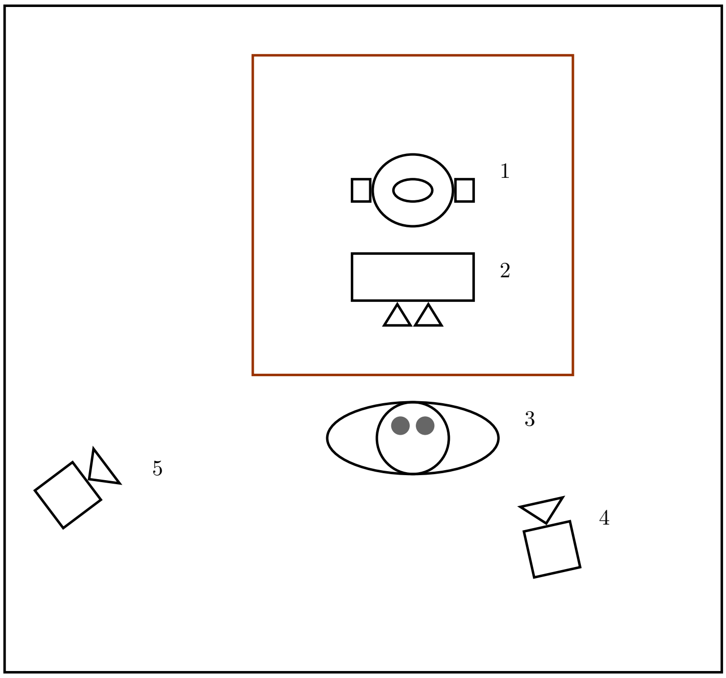

FIGURE 4.1: Illustration of the experiment setup seen from a top-down perspective. Scale not one-to-one. 1: Robot on table. 2: Eye-tracker on table. 3: Participant. 4: USB-camera for perspective view of gaze location (not present in virtual-presence test condition). 5: Video camera for experiment recording.

and one strong negative. The reason for this asymmetrical distribution is to prevent a false negative response, and is in line with the procedure of the source experiment in [20]. The participant is free to report any number of emotions for each ABM and can also express seeing no emotions. If the participants have not checked out a response it is regarded as a no. The second part regards the participants impression of motion speed and magnitude. It asks the participants to what degree they felt the speed of the expressed motion was fast or slow. The magnitude/speed items are graded on five-point Likert scales. The questionnaire is enclosed in appendix A.

A secondary questionnaire regarding the body parts on which the participant is focusing; head, arms, hands, upper body, legs and feet and 'other'. In case the eye-tracker malfunctions the participants can be asked to check the items and to what degree they paid attention to the corresponding body or motion parts. The secondary questionnaire is enclosed in appendix B. Each of the items has a two degree answer: 1. I Paid attention, 2. I paid no attention.

4.1.1 Participants

Two demographics groups of participants are recruited; seniors and young. The seniors are recruited in cooperation with SOSU Nord Future Lab, Ældre Sagen and Aktivitetscenteret Liselund. These co-operators ensured that all participants are "mentally fit", and thereby eliminating the need to apply screening procedures to identify participants suffering from dementia or depression prior to the experiment session. Another prerequisite for the senior participants is that their age ranged from 64 and upwards. The exclusion of participants suffering from dementia from experiment 1 is an adaptation of the requirements used in the study by Wong [32] and is explained in 2.1.3. Dementia is linked to impeded recognition of emotions. The upper boundary from Nomura's study is deemed arbitrary and neglected as age above 79 is not synonymous with cognitive decline. The young participants were recruited from classes of pedagogic assistant students at SOSU Nord with different ages ranging from 17 to 30 as the upper bound. All the participants were compensated with beverages, crisps, candy, coffee or cake after they had participated in the experiment.

4.1.2 Apparatus

The experiment use various equipment to conduct and document the experiment. For gaze tracking a Tobii X120 tracker was used to gather information on which areas of the robot the participant looked. To mediate the ABM a Nao H25 v3.3 robot by Alderbran Robotics was used. Its colors are white with grey plates. A Sony DV mini camera was used to record the experiment. A Logitech Pro 9000 webcam was used together with the Tobii Studios software version 3.2.1 for perspective recording of the participant's gaze on the robot. The layout of the equipment can be seen in figure 4.1. For the virtual condition a projector was used to show the virtual Nao.

4.1.3 Procedure

This section is a chronological review of how individual participants are guided through the experiment. Much care is taken to encourage the participant to ask questions if any should arise, that the participant is free to discontinue and leave the test at any time, that the test is focused on the system and not the person and finally that the test is anonymous.

The participant is explained that the experiment revolves around observing a robot. Before the experiment proceeds the participant is required to sign a consent form, that allows for the publication of the gathered data. Refreshments (coffee, cake, water) are offered. The participant is seated in front of a table with the eye-tracker and robot/projection in front of them as seen in figure 4.1. After being seated the eye-tracker is calibrated. Demographic regarding age is collected either before the calibration or before the session is concluded. The three types of robot ABM are expressed by the Robot in randomized order. After each ABM expression the participant is required to fill out the main questionnaire (Appendix A) regarding emotion/motion impression. After the experiment session ends the participant is asked if and what uses he/she could imagine to have of a robot. Interview question 1: "If you had a robot, and we imagined that this robot had the same capabilities as humans and could resemble anything. What should this robot do then?". Question 2: "What should it not do?". Finally the experiment and the participants is thanked.

4.2 Pilot Study of Experiment 1

A pilot study is designed and carry out before the first experiment is conducted, to evaluate the feasibility and duration of the experiment and to improve upon the experiment design. The pilot test uses the procedure described in section 4.1.3. Participants of the pilot study involves 9th semester Medialogy students as participants. These students allows for "new eyes" on the experiment design. The pilot study of experiment 1 uses a familiar setting. The Nao robot is placed on a desk in front of a chair as shown in figure 4.1. A camera is set up so it can record the participants reactions and document the experiment. The eye-tracking system was not enabled as technical issues arose, cause by the lighting conditions in the building.
Three Medialogy 9th semester students participated in the study. The pilot test confirmed the test procedure work. The importance of eliminating exogenous cues from the test setting was mentioned by all three participants.

4.3 Eye Tracking

The method used to measure gaze attention described in subsection 2.2.1, which uses questionnaires to measure which body parts the participants pays attention to while the robot expresses ABM. Attention paradigms [28] suggest that even if a person looks at a particular object, movement, event etc. doesn't mean that they are conscious about observing it. As a result the participants in Nomuras experiment may unconsciously pay attention to specific body parts of the robot, and not being aware of it. Therefore it is decided that using a eye-tracker will eliminate biased results from the use of questionnaires.

An eye-tracker is a device which tracks where the eyes gaze on a surface and provides x,y-coordinates as output, data that can be mapped as a 2D plot e.g. 4.3. The eye-tracker plots origin (0,0) in the upper left corner see figure 4.2. The eye-tracker software applies a filter to the raw gaze data and gives the fixation marks. Using the data acquired by the eye-tracker the problem of biased results from self-reported measures is eliminated, thereby providing more reliable data. The standard software package "Tobii Studio" is used and a backup of the data gathered in this project is located on the enclosed DVD.

4.3.1 Eye-tracker Issues

When analysing the fixation marks gathered from the eye-tracker, it was discovered that all the gathered fixation marks was not clustered directly on the robot, but had a slightly offset see figure 4.4. Since all exogenous cues are eliminated from the test setup's scene e.g. see image 4.35, it was ruled out as the cause for the fixation cluster. The robot having a initial start position were the clusters was also ruled out. Since great care as done to make sure the robot was centre of the tracking area and the it had the same initial start position throughout the whole of experiment 1. The cause for the offset, is suspected to be caused by a part of the eye-tracker setup, since the eye-trackers "camera perspective" was the only part of the setup which change position from test to test, which is consistent with that different offset in the four different tests. However since the eye-tracker was calibrated before each test by either virtual or using a calibration board, this should not be an issue. as it was the only part of the test setup which change position from test to test. Even through the fixation marks are not

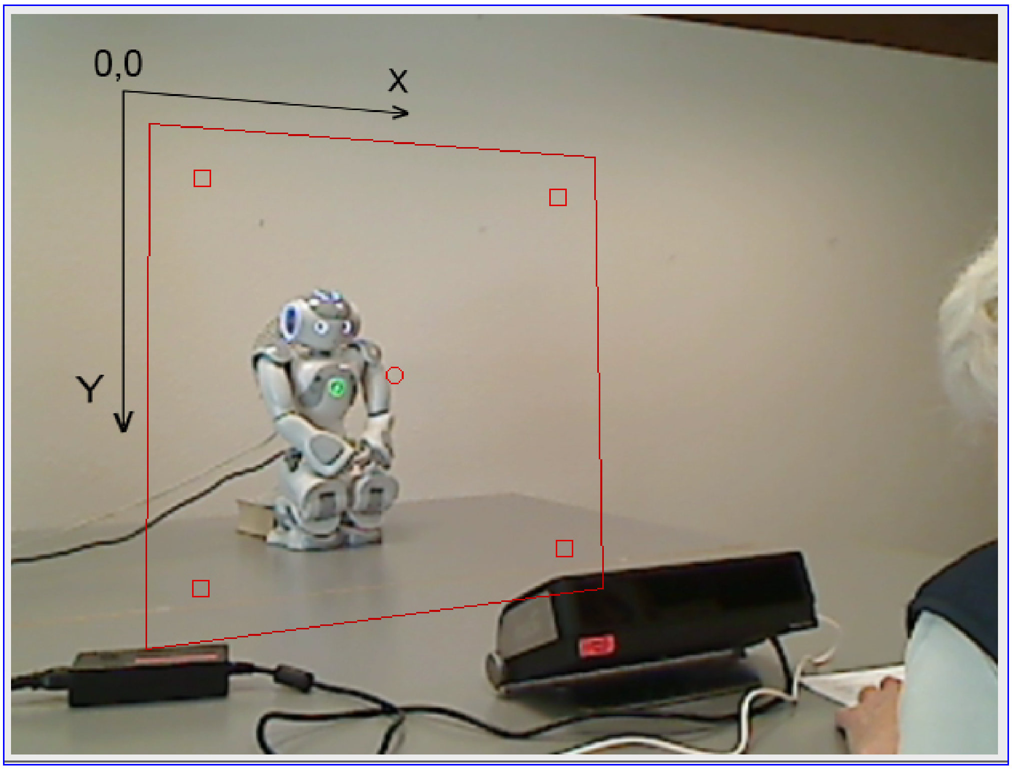

FIGURE 4.2: Screenshot from webcam recording "over the shoulder" for augmented reality preview of fixation movement. All physical-presence tests have this video recording. Feature of the Tobii setup. The black lines indicate how output values of fixation points are plotted with 0,0 in upper left corner.

directly within the region if the image where the robots, they still provide information were as the fixations are made in the upper region of the robot or lower region. Hence statistical procedures can, help determined if the fixation are made towards the top (x,0) or towards the bottom (x,480) of the image.

4.3.2 K-means

To locate which regions of the robot the participants look at, K-means cluster analysis is applied to the eye-tracker fixation coordinates. Furthermore K-means can obtain the number of fixations marks within each cluster, and thereby enable us to make a comparison between the different clusters. The K-means clustering algorithm is to used to classify or group objects based on the assumption that the data has attributes/features, which can classify or group objects together. The K-means clustering algorithm estimates the mean (vectors) of a set of groups, using the squared Euclidean distance from

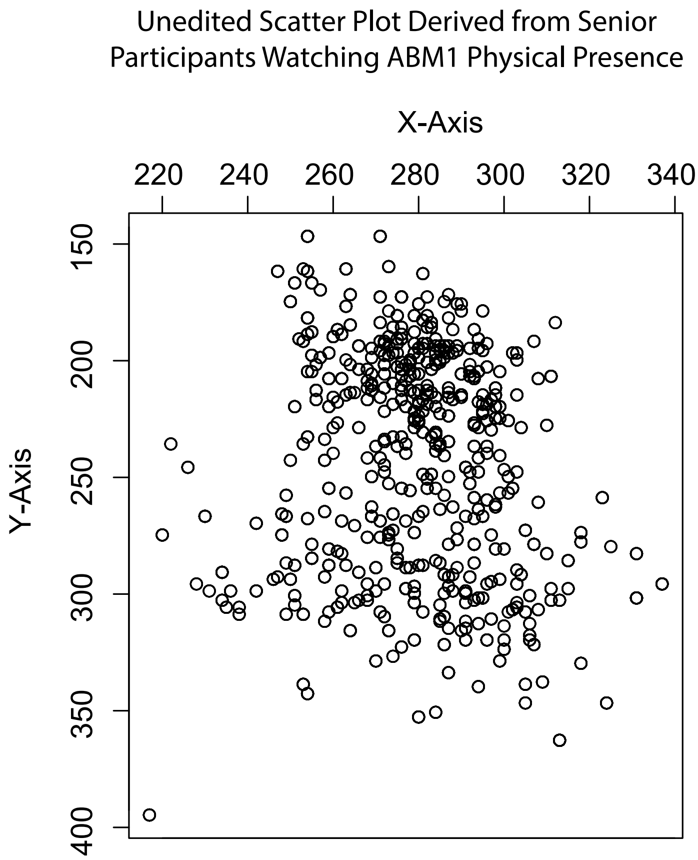

Figure 4.3: Unedited 2D plot from Vodskov Senior participants watching ABM1 - Anger

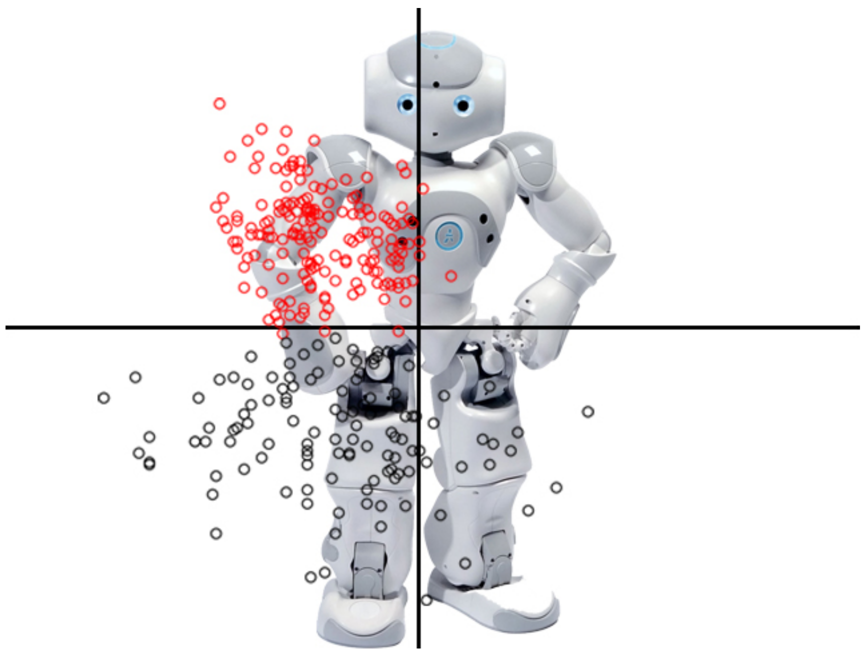

Figure 4.4: Image showing a example approximately where on the scene, the eye tracker ABM1 senior participants fixation marks are. Note that the robot is in the middle of the scene (recording resolution 640 width x 480 height), were the line meets is the centre of the scene.

each object (fixation) to each cluster and them computing them and assigning every object to the closest cluster means. The number of clusters are determined before running the K-means algorithm[25].

4.3.3 Cluster Analysis

This section briefly summaries some of the different statistical procedures, methods and software that has been used to identify which regions of the robot the participants paid attention to. An expert in statistics was consulted to verify which method was appropriate to use on the eye-tracker data set. Various methods like cluster analysis were discussed to find the best method of use. The open source software "RStudio", R Version 3.0.2 (2013-09-25) – "Frisbee Sailing" was used to analyse the eye-tracker fixation mark datasets, using the package "cluster" version 1.14.4.

4.3.4 K-means Procedure

Nomura et. al. [20] gathered information on specific parts of the robot the participants looked at, during his experiment. The initial reason for using the eye-tracker was to provide more precise information on which robot body parts the participants paid attention to. However the way the data is recorded from the eye-tracker makes it hard to do a precise comparison, between fixation marks and robot body parts. However using "C-Gap", gives a hint to where the participants are in fact looking at specific regions like i.e. arms, legs, head etc. as such each ABM fixation mark dataset is analysed with two types of K-means, one where the K-means algorithm looks for two means (the first referred to as "KM1", (and which uses the minimum requirement for running the K-means algorithm, which is looking for two means)) to examine if participants has regions of interest is in the upper region or lower region of the robot. The other K-means (is referred to as KM2) is intended to provide insight in which specific robot body parts the participants looks at, KM2 looks for the number of clusters suggested by the C-Gap method. Both KM1 and KM2 is analysed with a chi-square test to determine if there are differences between the clusters. If a difference is found in KM1 a binomial test is used to determine which cluster differ from the other. However because of the issues with the eye-tracker described in section 4.3.1, decided not apply other tests than chi-square tests. This decision was made because the offset described in section 4.3.1, made it impossible to superimpose the clusters directly on a image of a robot. However the chi-square test still gave insight in which ABMs the participants focus on specific parts and which they do not. The documentation for the R cluster package and fixation mark datasets is enclosed on the included DVD.

4.4 Physical Presence Eye-Tracker Results

This section presents the clusters and results of senior and young participants watching ABM's physical presence. There are 3 ABMs x 2 KMs x 2 presence conditions = a total of 12 KM analyses. A description of the procedure is in section 4.3.4.

Senior Physical Presence ABM1

KM1 yielded one cluster containing 120 fixation marks (cluster 1) with the x,y-means (287,291.6) and another cluster containing 183 fixation marks (cluster 2) with the x,y-means (283.2,205.6), presented in table 4.2 and plotted in figure 4.5. The chi-square test yielded a significant difference with df = 1 and a $p < 0.05$. The follow-up binominal test yielded a significant $p <$ 0,05 with a probability of success 60% that the participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of Fixation points
Cluster 1	287.0	291.	120
Cluster 2	283.2	205.6	183

Table 4.2: KM1 x,y-means and number of fixation marks of two clusters from senior participants in physical presence ABM1 - observing anger motion

KM2 yielded six clusters with various sizes and means presented in table 4.3 and plotted in figure 4.6. The chi-square test showed significantly one cluster differed from the other clusters with df = 5 and $p < 0.05$.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	284.5	187.8	72
Cluster 2	308.4	291.6	47
Cluster 3	260.4	285.5	43
Cluster 4	269.2	206.4	48
Cluster 5	292.1	230.2	75
Cluster 6	292.2	329.7	18

Table 4.3: KM2 x,y-means and number of fixation marks of six clusters from senior participants in Phycial Presence ABM1 - observing anger motion

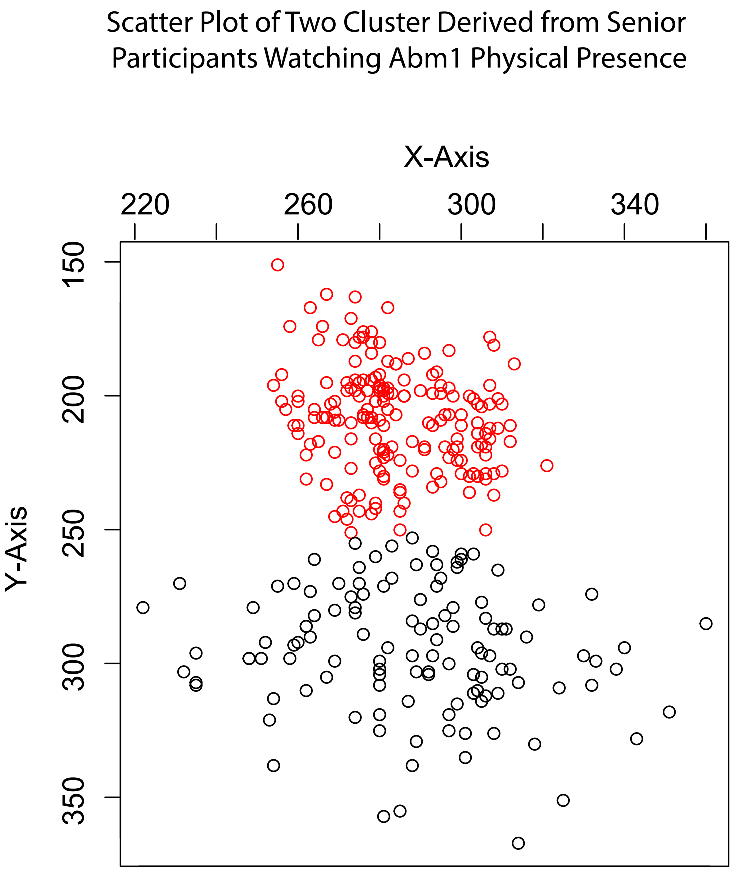

FIGURE 4.5: 2D data plot of six clusters from senior participant in physical presence ABM1 condition.

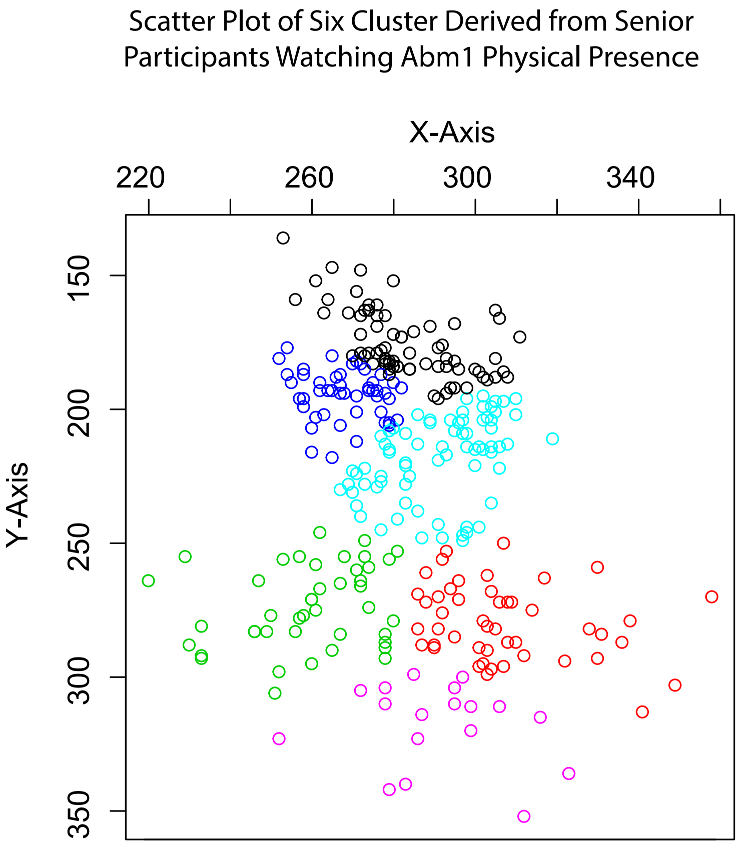

FIGURE 4.6: 2D data plot of six clusters from senior participant in physical presence ABM1 condition.

Young Physical Presence ABM1

KM1 yielded one cluster containing 78 fixation marks (cluster 1) with the x,y-means (320.96, 307.4) and another cluster containing 220 fixation marks (cluster 2) with the x,y-means (335.9, 204.2), presented in table 4.4 and plotted in figure 4.7. The chi-square test yielded a significant difference with df $= 1$ and a $p < 0.05$. The follow up binomial test yielded a significant $p < 0.05$, with a probability of success 73% that the participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	320.96	307.4	78
Cluster 2	335.9	204.2	220

TABLE 4.4: KM1 x,y-means and number of fixation marks of two clusters from young participants in physical presence ABM1 - observing anger motion

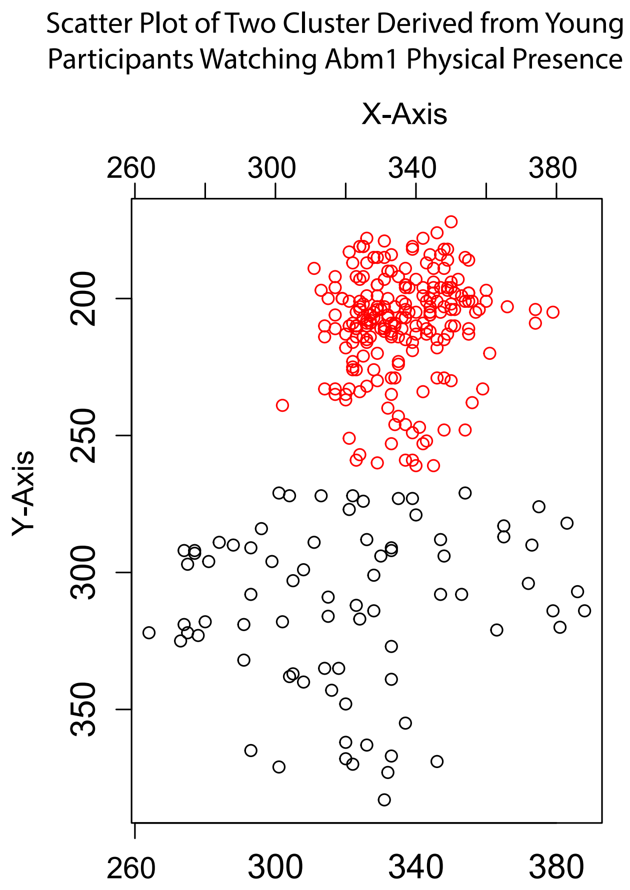

FIGURE 4.7: 2D data plot of two clusters from young participant in physical presence ABM1 condition.

KM2 yielded six clusters with various sizes and means presented in tablet 4.5 and plotted in figure 4.8. The chi-square test showed a significant difference df $= 5$, $p < 0.05$. suggesting that at leas one cluster significantly differed from the other clusters.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	320.6	349.4	20
Cluster 2	349.5	193.9	72
Cluster 3	292.1	298.9	29
Cluster 4	327.8	198.5	105
Cluster 5	331.9	239.04	48
Cluster 6	355.7	290.8	24

TABLE 4.5: KM2 x,y-means and number of fixation marks of six clusters from young participants in physical presence ABM1 - observing anger motion

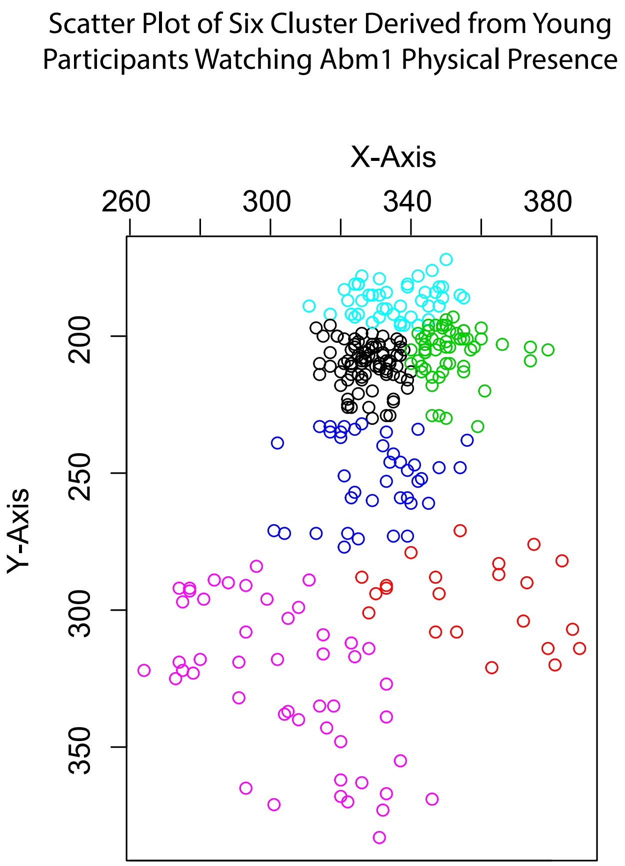

FIGURE 4.8: 2D data plot of six clusters from young participant in physical presence ABM1 condition.

4.4.1 Physical Presence ABM2

Section presenting the clusters of senior and young participants watching ABM2 physical presence.

Senior Physical Presence ABM2

KM1 yielded one cluster containing 105 fixation marks (cluster 1) with the x,y-means (292.1, 285.7) and another cluster containing 278 fixation marks (cluster 2) with the x,y-means (289.2, 201.4), presented in table 4.6 and plotted in 4.9. The chi-square test yielded a significant difference df = 1, $p < 0.05$. The follow up binomial test yielded a significant $p < 0.05$, with a probability of success 72% that the participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	292.1	285.7	105
Cluster 2	289.2	201.4	278

Table 4.6: KM1 x,y-means of two clusters from senior participants in physical presence ABM2 - observing happiness motion.

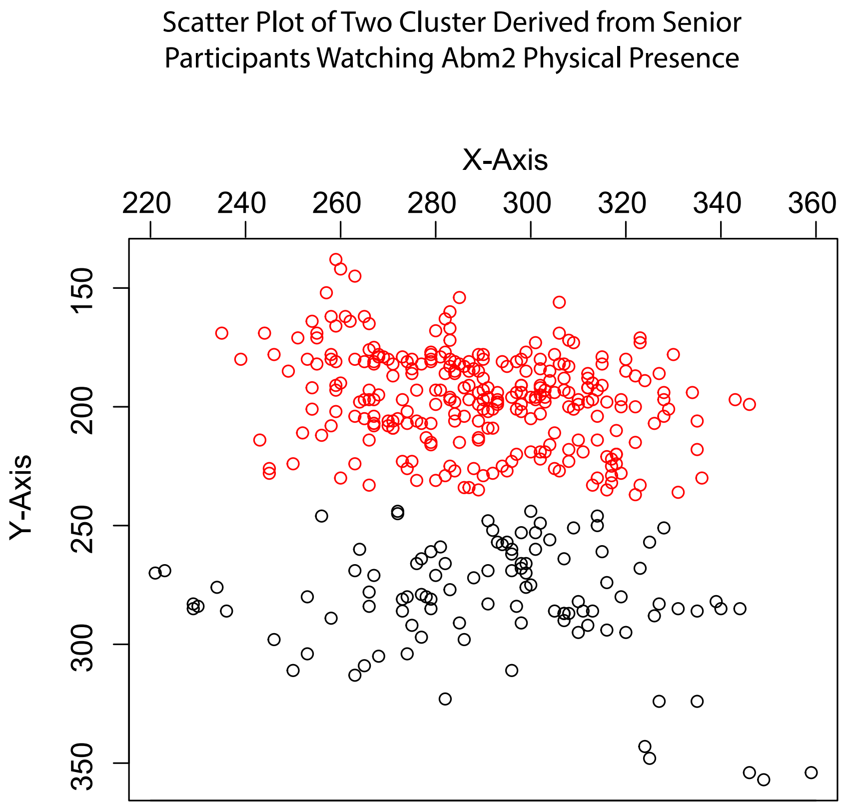

Figure 4.9: 2D data plot of six clusters from senior participant in physical presence ABM2 - observing happiness.

KM2 yielded three clusters with various sizes and means presented in table 4.7 and plotted in figure 4.10. The chi-square test showed significantly one cluster differed from the others with df = 2 and $p < 0.05$.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	270.3	194.03	133
Cluster 2	292.1	285.7	105
Cluster 3	306.5	208.1	145

TABLE 4.7: KM2 x,y-means and number of fixation coordinates of three clusters from young participants in physical presence ABM2 - observing happiness motion

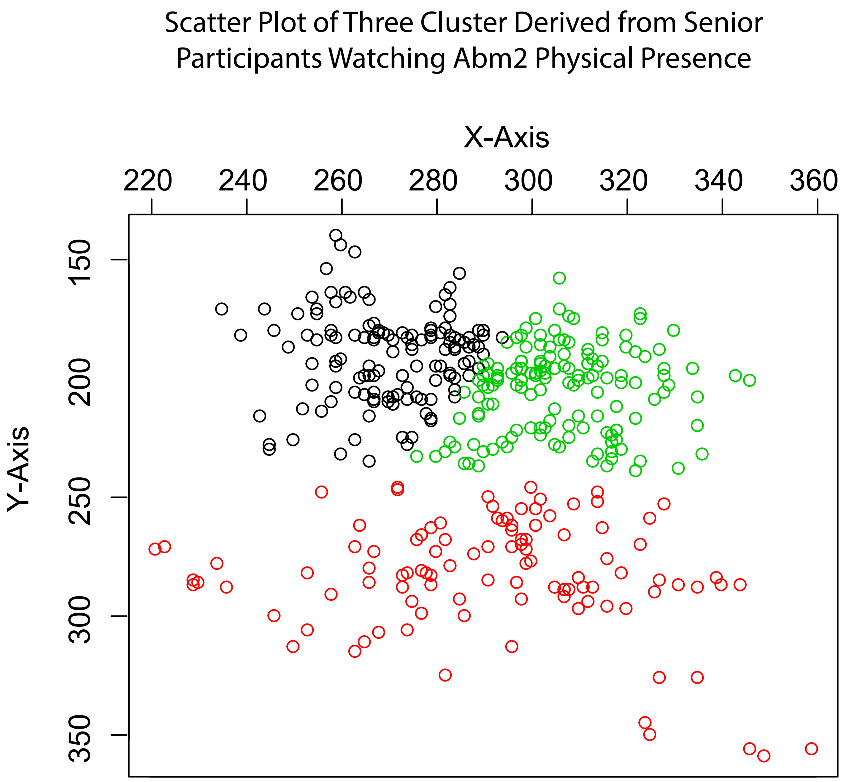

FIGURE 4.10: 2D data plot of six clusters from senior participant in physical presence ABM2 - observing happiness.

Young Physical Presence ABM2

KM1 yielded one cluster containing 100 fixation marks (cluster 1) with the x,y-means (328.98, 277.4) and another cluster containing 182 fixation marks (cluster 2) with the x,y-means (335.6, 194.6), presented in table 4.8 and plotted in 4.11. The chi-square test yielded a significant difference df = 1, $p < 0.05$. The follow up binomial test yielded a significant $p < 0.05$, with a probability of success 64% that the participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	328.98	277.4	100
Cluster 2	335.6	194.6	182

TABLE 4.8: KM1 x,y-means of two clusters from young participants in physical presence ABM2 - observing happiness motion

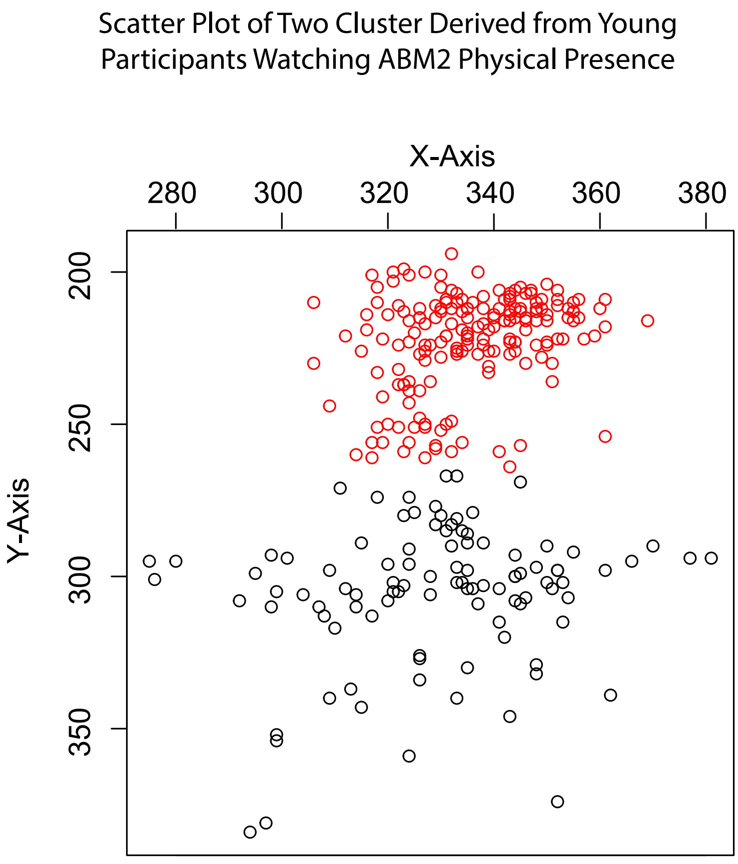

FIGURE 4.11: 2D data plot of two clusters from young participant in physical presence ABM2 - observing happiness.

KM2 yielded seven clusters with various sizes and means presented in table 4.9 and plotted in figure 4.12. The chi-square test showed significantly one cluster differed from the others with df = 6, $p < 0.05$.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	307.4	274.7	27
Cluster 2	344.8	271.1	42
Cluster 3	324.9	317.9	18
Cluster 4	328.4	234.02	38
Cluster 5	327.3	200.1	45
Cluster 6	347.7	187.4	72
Cluster 7	327.6	180.9	40

TABLE 4.9: KM2 x,y-means and number of fixation marks of seven clusters from young participants in physical presence ABM2 - observing happiness motion

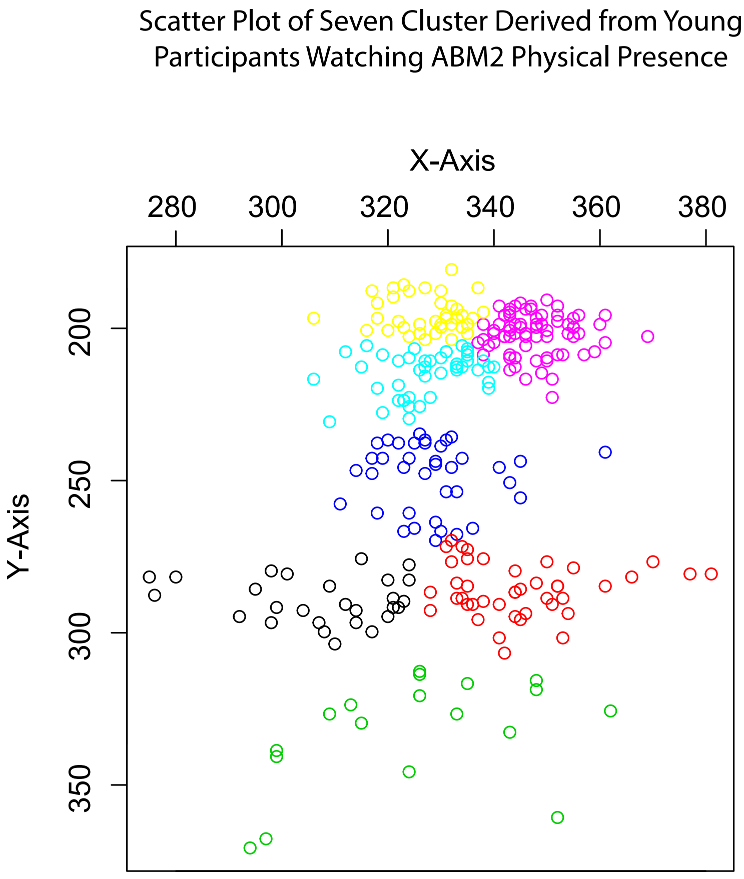

FIGURE 4.12: 2D data plot of seven clusters from young participant in physical presence ABM2 - observing happiness.

Senior Physical Presence ABM3

KM1 yielded one cluster containing 96 fixation marks (cluster 1) with the x,y-means (288.8, 281.8) and another cluster containing 240 fixation marks (cluster 2) with the x,y-means (284.8, 199.8), presented in table 4.10 and plotted in 4.13. The chi-square test yielded a significant difference df = 1, $p < 0.05$. The follow up binomial test yielded a significant $p < 0.05$, with a probability of success 71% that the participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	288.8	281.8	96
Cluster 2	284.8	199.8	240

TABLE 4.10: KM1 x,y-means of two clusters from senior participants in physical presence ABM3 - observing sadness motion

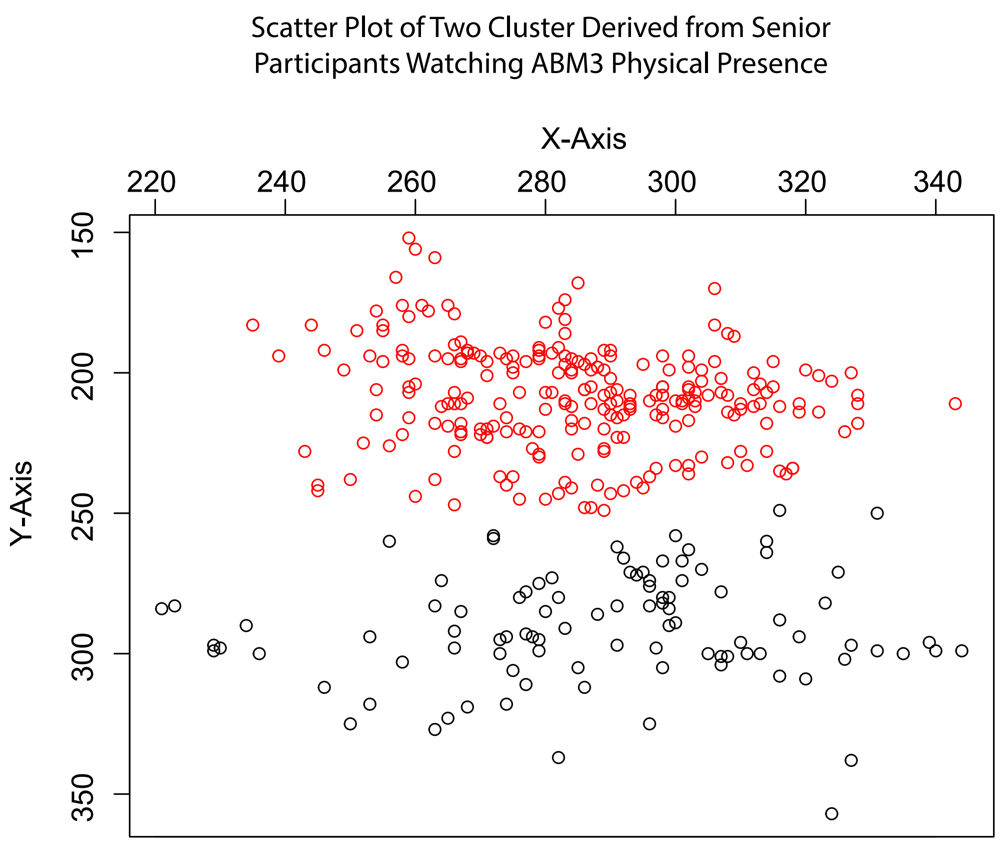

FIGURE 4.13: 2D data plot of two clusters from senior participant in physical presence ABM3 - observing sadness.

KM2 yielded seven clusters with various sizes and means presented in table 4.11 and plotted in figure 4.14. The chi-square test showed significantly one cluster differed from the others with df = 6, $p < 0.05$.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	318.3	295.1	24
Cluster 2	294.7	210.2	66
Cluster 3	311.3	195.3	49
Cluster 4	269.3	181.1	75
Cluster 5	294.95	262.7	41
Cluster 6	267.5	217.9	48
Cluster 7	260.5	292.9	33

TABLE 4.11: KM2 x,y-means and number of fixation marks of seven clusters from young participants in physical presence ABM3 - observing sadness motion

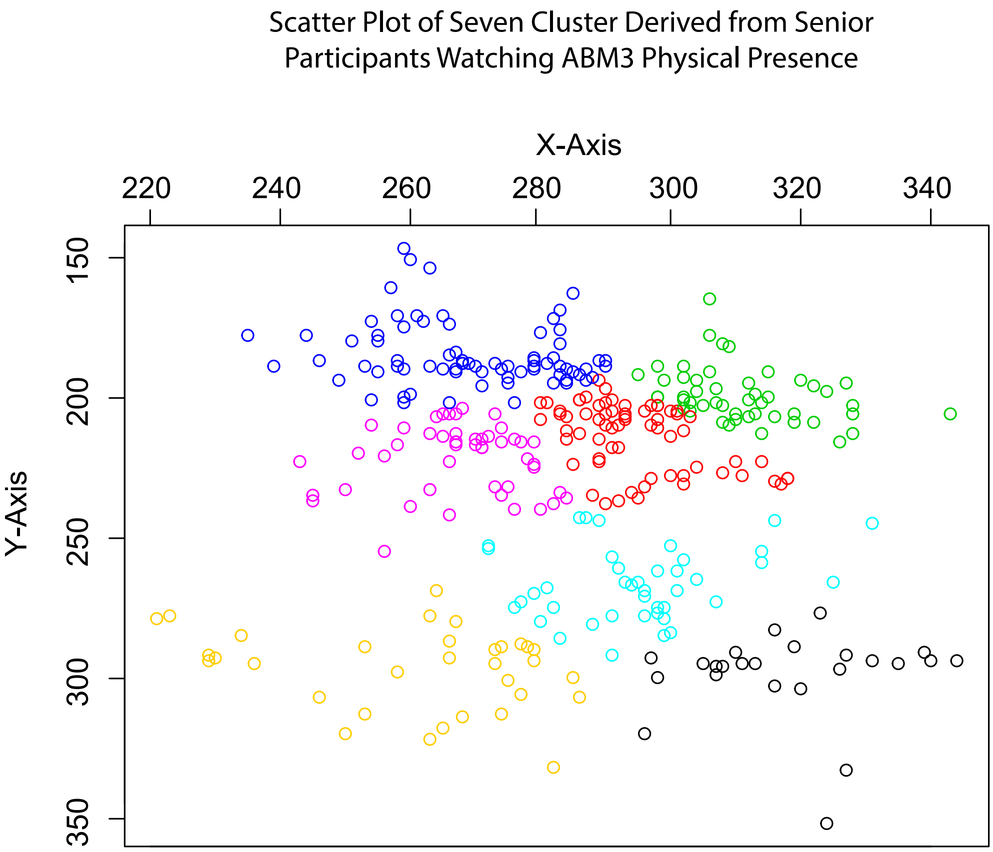

FIGURE 4.14: 2D data plot of seven clusters from senior participant in physical presence ABM3 - observing sadness.

Young Physical Presence ABM3

KM1 yielded one cluster containing 230 fixation marks (cluster 1) with the x,y-means (325.7, 207.9) and another cluster containing 103 fixation marks (cluster 2) with the x,y-means (329.9, 301.4), presented in table 4.12 and plotted in 4.15. The chi-square test yielded a significant difference df = 1, $p < 0.05$. The follow up binomial test yielded a significant $p < 0.05$, with a probability of success 75% that the participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	325.7	207.9	230
Cluster 2	329.7	301.4	103

TABLE 4.12: KM1 x,y-means of two clusters from young participants in physical presence ABM3 - observing sadness motion

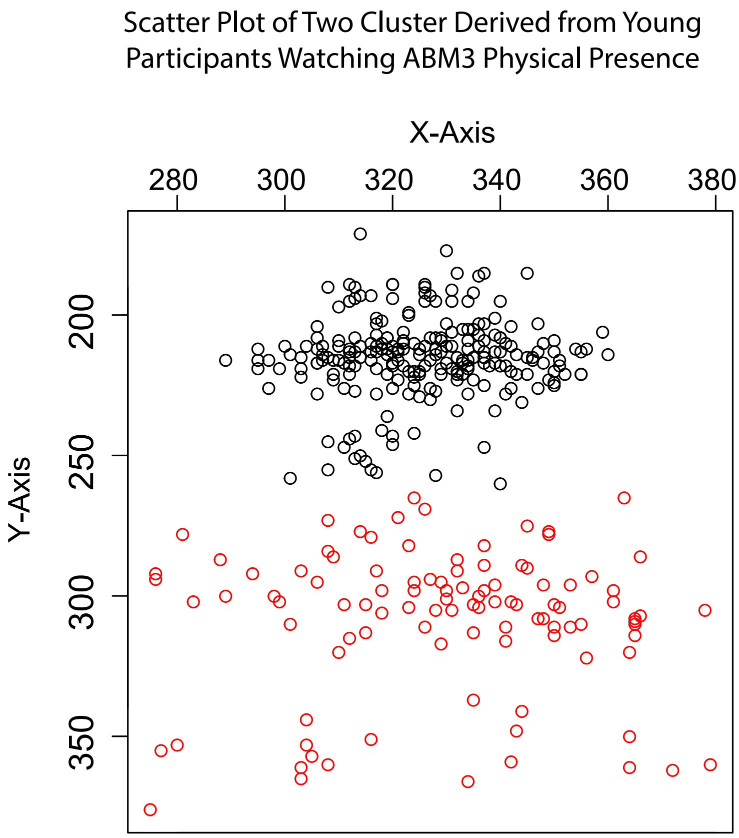

FIGURE 4.15: 2D data plot of two clusters from young participant in physical presence ABM3 - observing sadness.

KM2 yielded seven clusters with various sizes and means presented in table 4.13 and plotted in figure 4.16. The chi-square test showed significantly one cluster differed from the others with df = 3, $p < 0.05$.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	323.8	348.7	19
Cluster 2	331.3	289.8	86
Cluster 3	314.2	212.1	114
Cluster 4	337.02	203.0	114

TABLE 4.13: KM1 x,y-means of four clusters from young participants in physical presence ABM3 - observing sadness motion

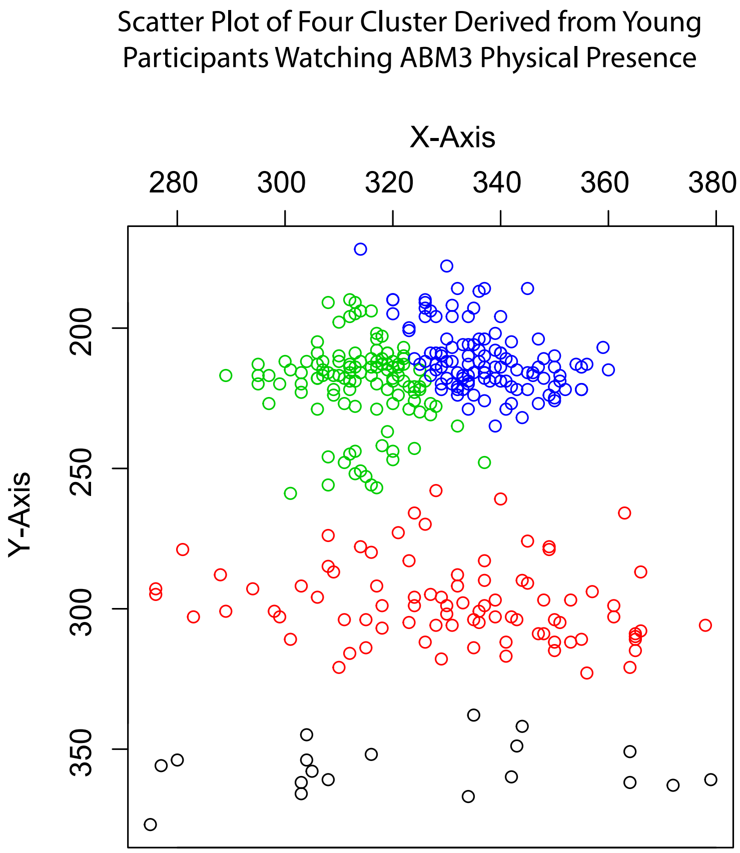

FIGURE 4.16: 2D data plot of four clusters from young participant in physical presence ABM3 - observing sadness.

4.4.2 Virtual Presence Eye-Tracker Results

Section presenting the clusters and results of senior and young participants watching ABM's virtual presence.

Senior Virtual Presence ABM1

KM1 yielded one cluster containing containing 74 fixation marks (cluster 1) and another cluster containing 116 fixation marks (cluster 2), with cluster 1's x,y-cluster means (525.7,480.5) and cluster 2's x,y- cluster means (548.5,289.1), see tale 4.14 and figure

4.17 for overview. The follow up chi-square test yielded a significant difference with df = 1 and p-value = 0.002311. The binomial test yielded significant p-value = 0.002337 and a probability of success of 61% that the participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	525.7	480.5	74
Cluster 2	548.5	289.1	116

TABLE 4.14: KM1 x,y-means and number of fixation marks of two clusters from senior participants in virtual presence ABM1 - observing anger motion

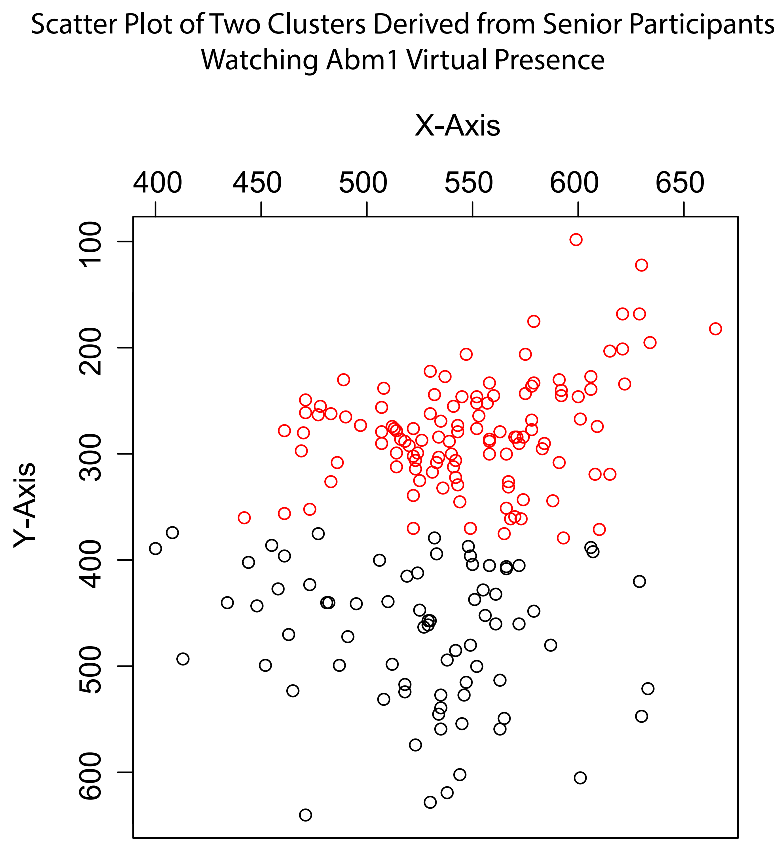

FIGURE 4.17: 2D data plot of two clusters from senior participant in virtual presence ABM1 condition.

KM2 yielded seven clusters presented with various sizes and means presented in table 4.15 and plotted in figure 4.18. The chi-square test showed that there was no significant difference, with df = 6, $p > 0.05$.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	457.9	428.0	21
Cluster 2	531.4	308.7	36
Cluster 3	621.4	178.0	9
Cluster 4	539.8	556.5	26
Cluster 5	487.6	277.3	17
Cluster 6	572.4	261.9	34
Cluster 7	551.1	446.6	26
Cluster 8	574.4	366.1	21

TABLE 4.15: KM2 x,y-means and number of fixation coordinates of Eight clusters from senior participants in virtual presence ABM1 - observing anger motion

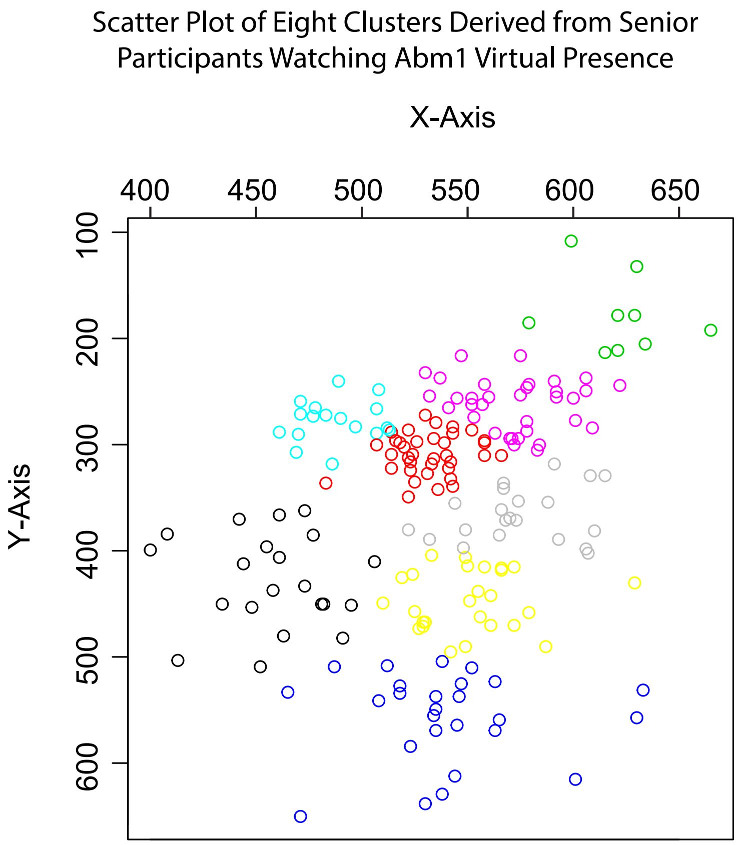

FIGURE 4.18: 2D data plot of Eight clusters from senior participant in virtual presence ABM1 condition.

Young Virtual Presence ABM1

KM1 yielded one cluster containing 98 fixations marks (cluster 1) and another cluster containing 67 fixation marks (cluster 2), with cluster 1's x,y-clusters means (432.1, 231.8) and cluster 2's x,y- cluster means (421.4, 382.3), see table 4.16 and figure 4.19 for overview. The follow up chi-square test yielded a significant difference with df = 1 and $p < 0.05$. The binomial test yielded significant $p < 0.05$ and a probability of success of 59% that the participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	432.1	231.8	98
Cluster 2	421.4	382.3	67

TABLE 4.16: KM1 x,y-means and number of fixation marks of two clusters from young participants in virtual presence ABM1 - observing anger motion

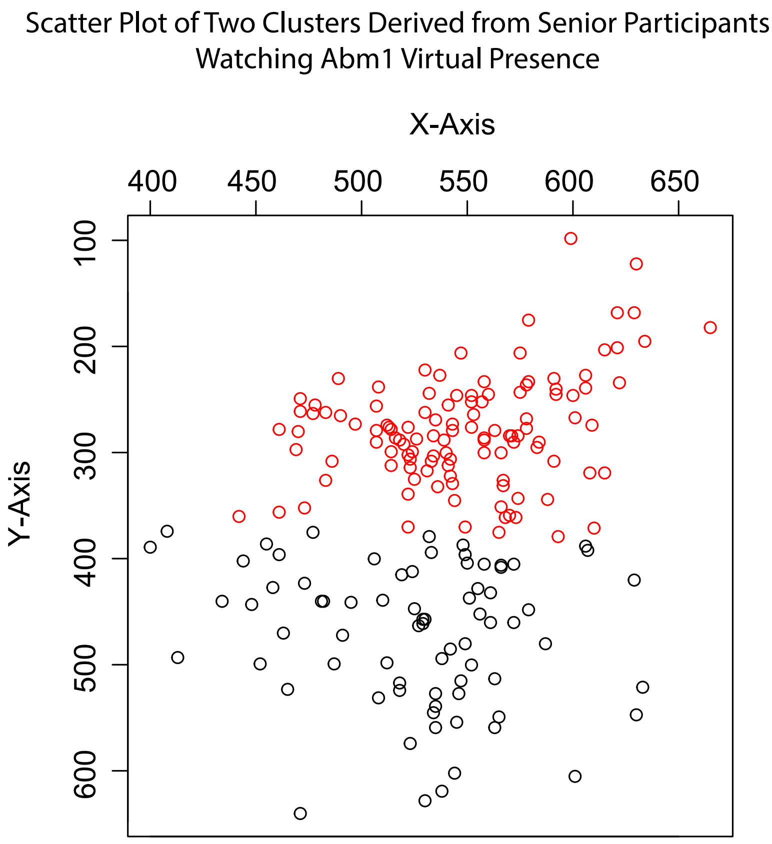

FIGURE 4.19: 2D data plot of two clusters from young participant in virtual presence ABM1 condition.

KM2 yielded eight clusters presented with various sizes and means presented in table 4.17 and plotted in figure 4.20. The chi-square test showed that at least one differed significantly, with df = 7 and $p < 0.05$.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	405.8	501.2	5
Cluster 2	415.1	356.0	19
Cluster 3	455.1	314.8	17
Cluster 4	391.6	325.9	8
Cluster 5	433.9	261.4	30
Cluster 6	407.3	433.6	12
Cluster 7	435.9	397.9	14
Cluster 8	429.7	208.1	60

TABLE 4.17: KM2 x,y-means and number of fixation coordinates of Eight clusters from young participants in virtual presence ABM1 - observing anger motion

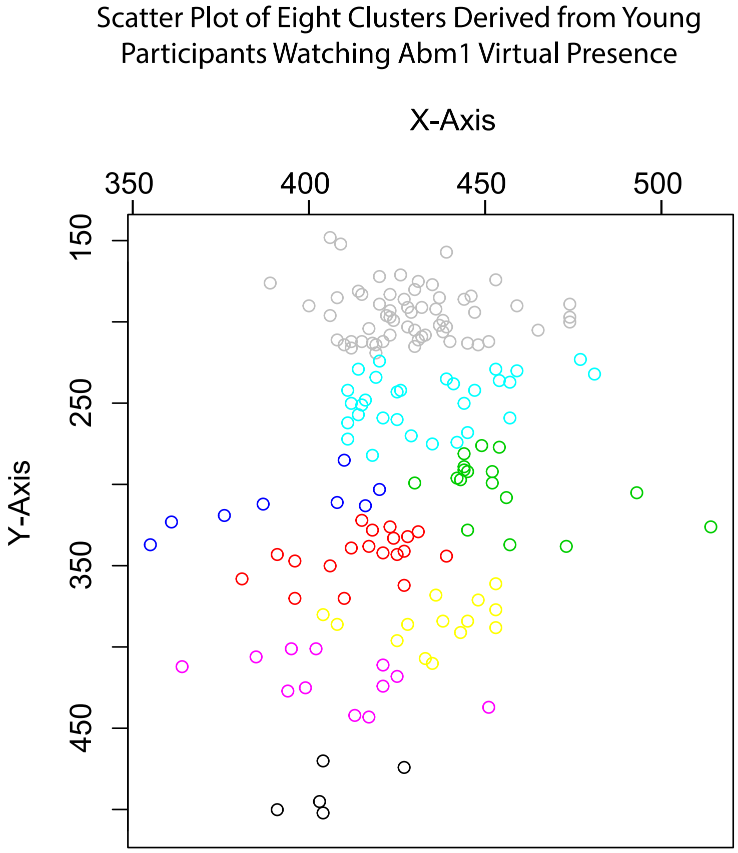

FIGURE 4.20: 2D data plot of eight clusters from young participant in virtual presence ABM1 condition.

Virtual Presence ABM2 Section presenting the clusters of senior and young participants watching ABM2 virtual presence.

4.4.2.1 Virtual Presence Senior ABM2

KM1 yielded one cluster contaning 65 fixation marks (cluster 1) and another cluster containing 118 fixation marks (cluster 2), with clusters 1's x,y-clusters means (546.6, 503.7) and cluster 2's x,y-cluster means (568.2, 264.2), see table 4.18and figure 4.21 for overview. The follow up chi-square test yielded a significant result, with df = 1, and $p < 0.05$. The follow up binominal test yielded a $p < 0.05$ and a probability of success of 64% that the participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	546.6	503.7	65
Cluster 2	568.2	264.2	118

Table 4.18: KM1 x,y-means and number of fixation marks of two clusters from senior participants in virtual presence ABM2 - observing happiness motion

KM2 yielded seven clusters presented with various sizes and means presented in table4.19 and plotted in figure 4.22. The chi-square test showed that none of the clusters differed significant form the other clusters, with df = 6 and $p > 0.05$.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	531.3	201.5	21
Cluster 2	552.8	581.3	26
Cluster 3	542.4	470.0	29
Cluster 4	536.0	289.4	30
Cluster 5	635.7	168.5	15
Cluster 6	552.3	368.7	33
Cluster 7	602.5	260.8	29

Table 4.19: KM2 x,y-means and number of fixation coordinates of seven clusters from senior participants in virtual presence ABM2 - observing happiness motion

Virtual Presence Young ABM2

KM1 yielded one cluster containing 68 fixation marks (cluster 1) and another cluster containing 89 fixation marks (cluster 2), with cluster 1's x,y -cluster means (432.2, 355.4) and cluster 2's x,y-cluster means (436.2, 210.6), plottede in figure 4.20see table 4.20 and

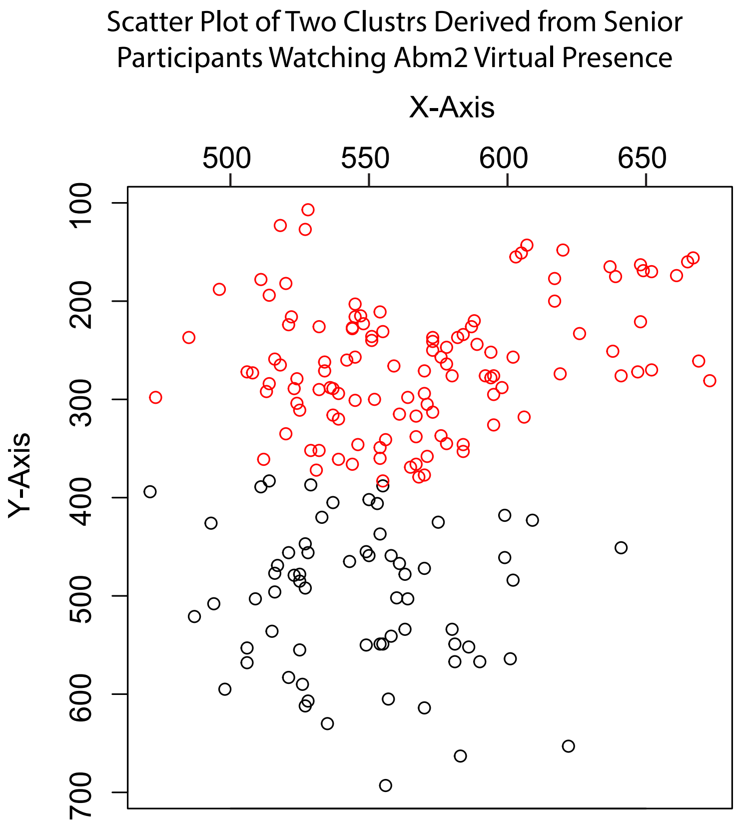

FIGURE 4.21: 2D data plot of two clusters from senior participant in virtual presence ABM2 condition.

for overview. The follow up chi-square test did not yield a significant result, with df = 1 and $p > 0.05$.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	432.2	355.4	68
Cluster 2	436.2	210.6	89

TABLE 4.20: KM1 x,y-means and number of fixation marks of two clusters from young participants in virtual presence ABM2 - observing happiness motion

KM2 yielded seven clusters presented with various sizes and means presented in table 4.21 and plotted in figure 4.24. The chi-square test showed that at least one differed significantly, with df = 6 and $p < 0.05$.

Virtual Presence Senior ABM3

Section presenting the clusters of senior and young participants watching ABM3 virtual presence.

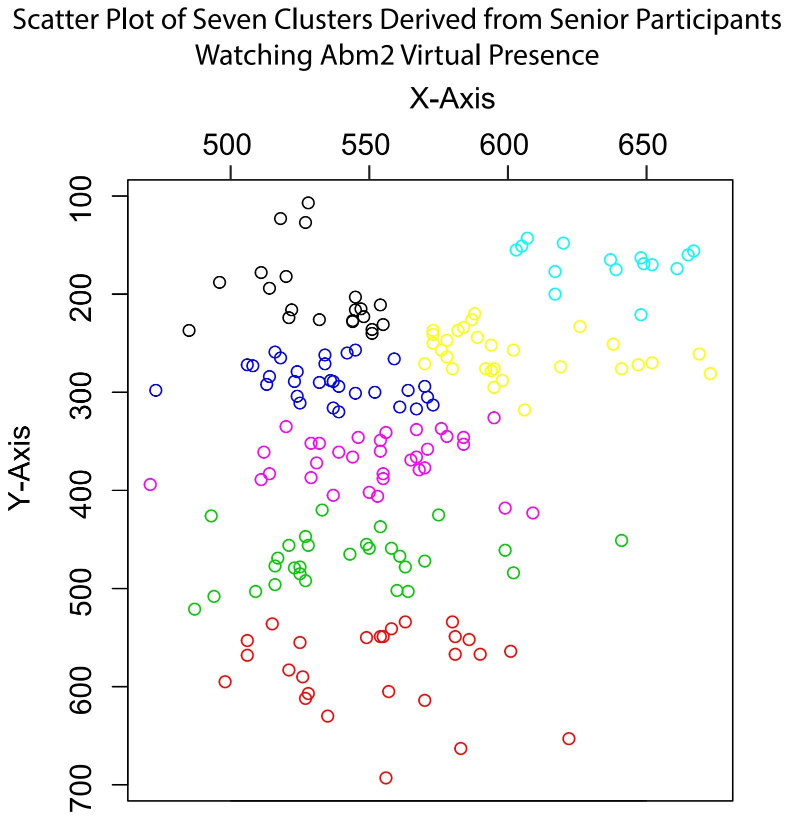

FIGURE 4.22: 2D data plot of seven clusters from senior participant in virtual presence ABM2 condition.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	413.7	193.4	25
Cluster 2	452.2	204.8	23
Cluster 3	425.0	341.3	29
Cluster 4	428.0	413.5	22
Cluster 5	444.3	156.4	12
Cluster 6	452.6	296.7	23
Cluster 7	434.3	246.4	23

TABLE 4.21: KM2 x,y-means and number of fixation coordinates of seven clusters from young participants in virtual presence ABM2 - observing happiness motion

KM1 yielded one cluster containing 143 fixation marks (cluster 1) with x,y-cluster means(547.7,301.1) and another cluster containing 68 fixation marks (cluster 2) with x,y-cluster means (529.2,483.6), plotted in figure 4.25 and displayed in table 4.22. The follow up chi-square test yielded a significant difference df = 1, $p < 0.05$. The binominal test yield a $p < 0.05$ and a probability of success of 67% that the participant looked at the upper region of the robot.

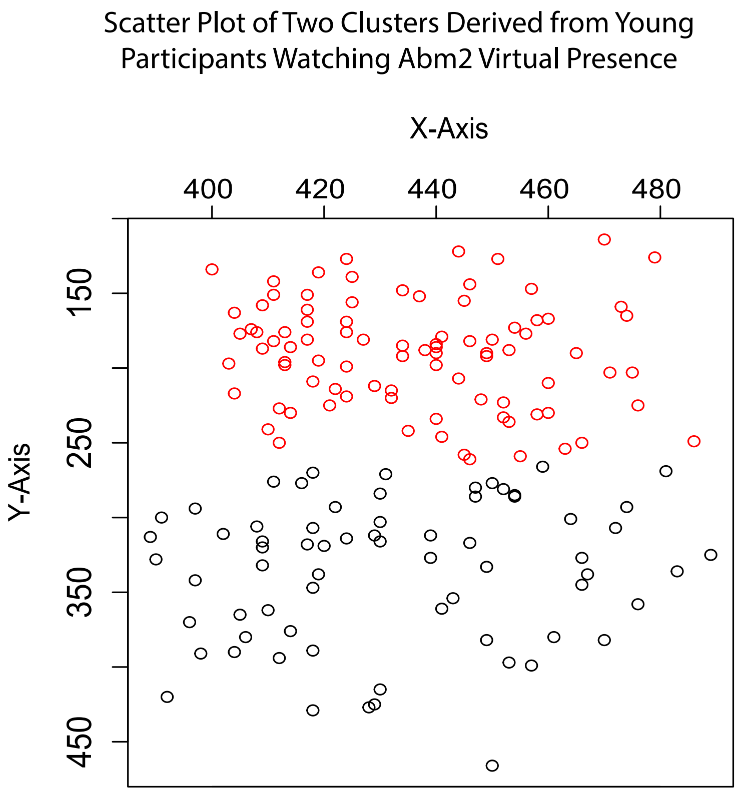

FIGURE 4.23: 2D data plot of eight clusters from young participant in virtual presence ABM2 condition

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	547.7	301.1	143
Cluster 2	529.2	483.6	68

TABLE 4.22: KM1 x,y-means and number of fixation marks of two clusters from senior participants in virtual presence ABM3 - observing sadness motion

KM3 yielded seven clusters presented with various sizes and means presented in table 4.23 and plotted in figure 4.26. The chi-square test showed that at least one cluster differed significantly, with df = 6, $p < 0.05$.

Virtual Presence Young ABM3

KM1 yielded one cluster containing 136 fixation marks (cluster 1) and another cluster containing 77 fixation marks (cluster 2), with cluster 1's x,y-cluster means (422.02, 223.1) and cluster 2's x,y-cluster means (423.8, 366.5), plotted in 4.27 see table 4.24 for overview. The follow up chi-square test yielded a significant difference with df = 1 and $p < 0.05$. The binominal test yielded a $p < 0.05$ and a probability of success 63% that

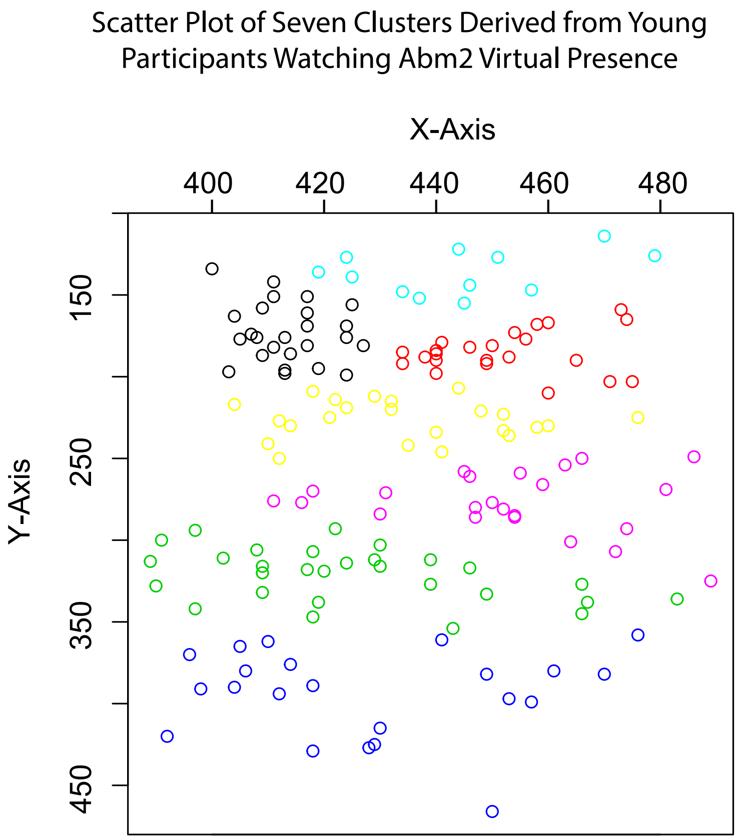

FIGURE 4.24: 2D data plot of seven clusters from young participant in virtual presence ABM2 condition.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	493.2	284.9	20
Cluster 2	524.8	396.0	26
Cluster 3	582.8	315.8	37
Cluster 4	583.2	479.2	20
Cluster 5	534.9	314.7	53
Cluster 6	500.2	517.8	33
Cluster 7	577.8	223.2	22

TABLE 4.23: KM2 x,y-means and number of fixation coordinates of seven clusters from seven participants in virtual presence ABM2 - observing sadness motion

the participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	422.0	223.1	136
Cluster 2	423.8	366.5	77

TABLE 4.24: KM1 x,y-means and number of fixation marks of two clusters from young participants in virtual presence ABM3 - observing sadness motion

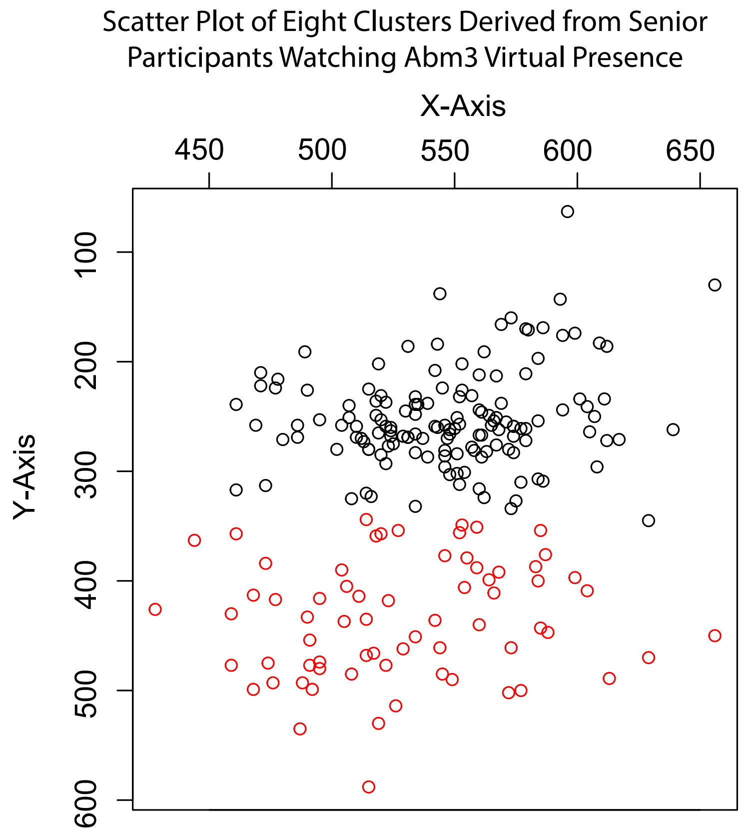

FIGURE 4.25: 2D data plot of two clusters from senior participant in virtual presence ABM3 condition.

KM2 yielded ten clusters presented with various sizes and means presented in table 4.25 and plottede in figure 4.28. The chi-square test showed that at least one differed significantly, with df = 9 and $p < 0.05$.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation marks
Cluster 1	435.9	380.4	22
Cluster 2	451.2	246.6	15
Cluster 3	420.8	225.5	36
Cluster 4	410.1	272.1	26
Cluster 5	386.0	202.5	11
Cluster 6	382.6	415.7	15
Cluster 7	443.9	205.2	19
Cluster 8	418.4	183.6	29
Cluster 9	484.6	362.6	14
Cluster 10	404.7	328.5	26

TABLE 4.25: KM2 x,y-means and number of fixation coordinates of seven clusters from seven participants in virtual presence ABM3 - observing sadness motion

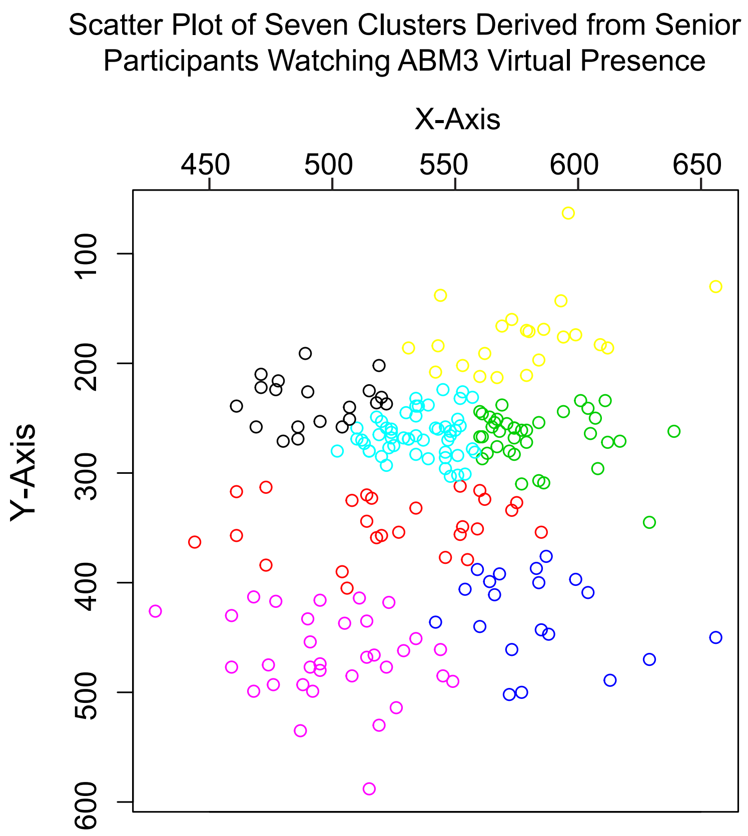

FIGURE 4.26: 2D data plot of seven clusters from senior participant in virtual presence ABM3 condition.

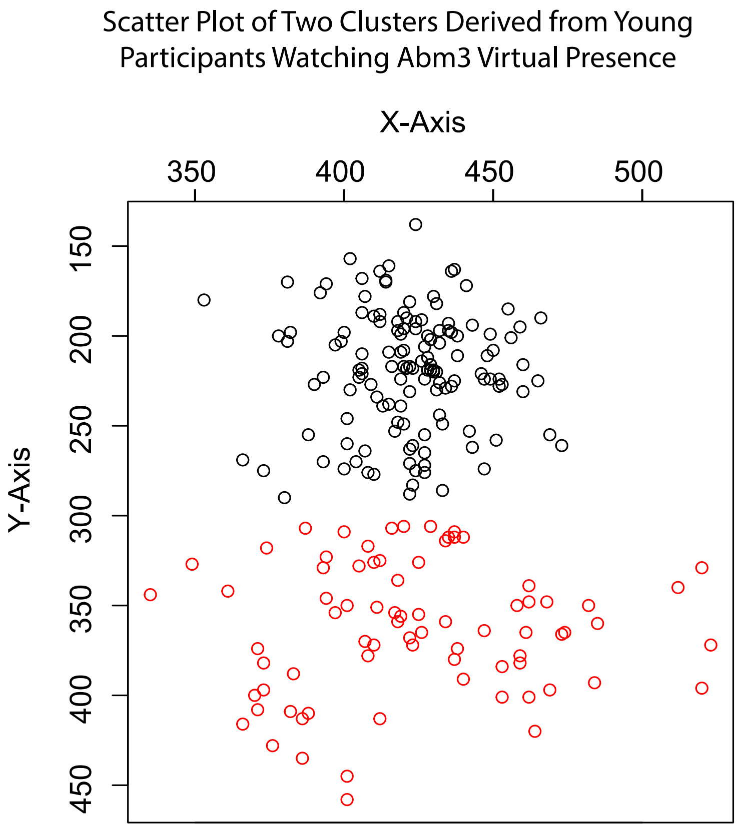

FIGURE 4.27: 2D data plot of two clusters from young participant in virtual presence ABM3 condition.

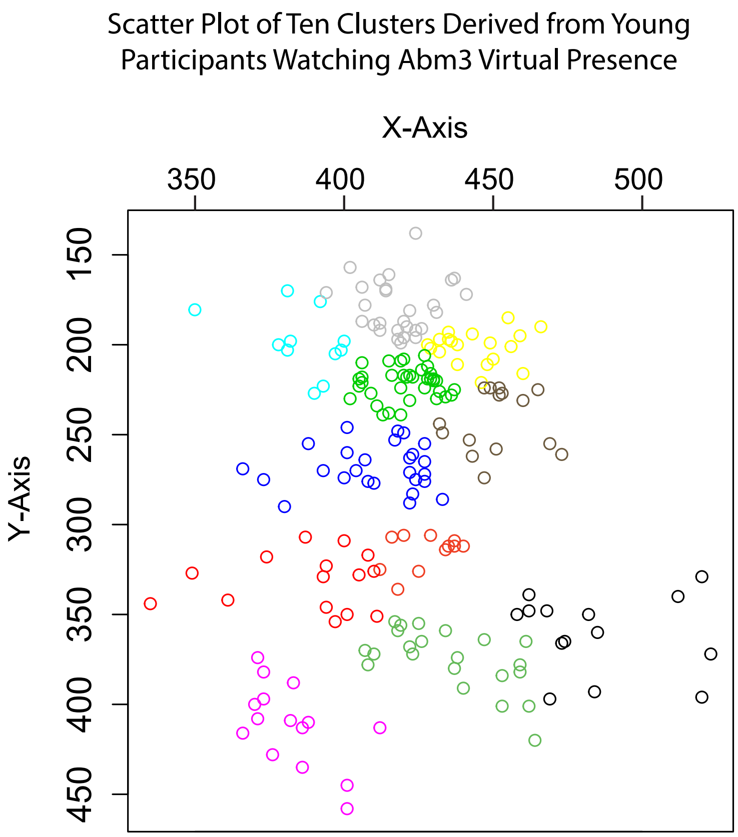

FIGURE 4.28: 2D data plot of ten clusters from young participant in virtual presence ABM3 condition.

4.5 Simplified Fixation Results

In addition to cluster analysis of fixation patterns an analysis is also performed on the vertical axis of fixation data by isolating the y-axis. A series of independent t-tests are used to compare y-axis fixations between students and seniors by compacting the 2D measurements to a single dimension. The purpose of this comparison is to test if a significant difference exists in fixations along the y-axis. The "cognitive bias" referred to by Nomura [20] is that seniors fixate lower on the robot than students when viewing sadness and higher on the robot when viewing anger. Both physical and virtual presence is covered in this section and it should be noted that figures show fixation points inverted. Since the origo of the original data is recorded in the top left corner of the coordinate system data is plotted upside-down. N is the number of fixations per group and M is the mean value along the y-axis (lower towards the robot's head). Effect size 0.1 is a small effect, 0.3 is medium and 0.5 is large.
Tests are run on fixation points along the y-axis and are summarized for all students and seniors and all calculations are placed on the enclosed DVD in folder 'Experiment 1\spss output fixation y axis analysis'. For physical presence a significant difference is found for ABM 1 and 3: When displaying anger y-axis coordinates for students (N = 298, M = 231.3, Std = 50.6) on average looked up higher on the robot than seniors (N = 303, M = 239.6, Std = 47.2) with results t(599) = -2,101, $p < 0.05$, effect size = 0.086. When displaying sadness y-axis coordinates for seniors (N = 336, M = 223,2, Std = 41.6) on average looked up higher on the robot than students (N = 333, M = 236.8, Std = 47.2) with test results t(655.4) = 3.957, $p < 0.05$, effect size = 0.153. A plot of the fixations of ABM3 is shown in figure 4.29.

Results of independent t-tests in the virtual presence condition shows significant difference between students and seniors for all 3 ABMs.
ABM1: Students looked high up on the robot (N = 165, M = 292.9, Std = 85,7) and the seniors had mean fixation point (N = 190, M = 363.6, Std = 110.2), t(348.9) = -6.676, $p < 0.05$, effect size = 0.337
ABM2: Students looked high up on the robot (N = 157, M = 273.3, Std = 82.99) compared to seniors (N = 183, M = 349.3, Std = 133.9), t(309.1) = -6.377 ,$p < 0.05$, effect size = 0.341.
ABM3: Students looked high up on the robot (N = 213, M = 274.9, Std = 77.2), seniors

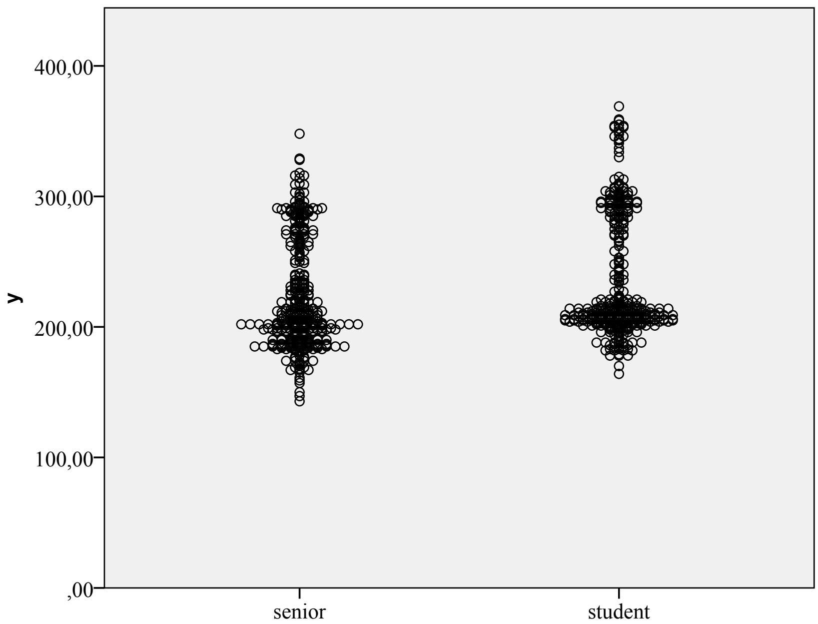

FIGURE 4.29: Fixations along the y-axis from ABM3 in the physical presence condition. Seniors fixate slightly higher on the robot when displaying sadness. If x-values are identical the points stack horizontally.

(N = 211, M = 359.9, Std = 97.9), t(398.4) = -9.912, $p < 0.05$, effect size = 0.445. See figure 4.30 for plots.

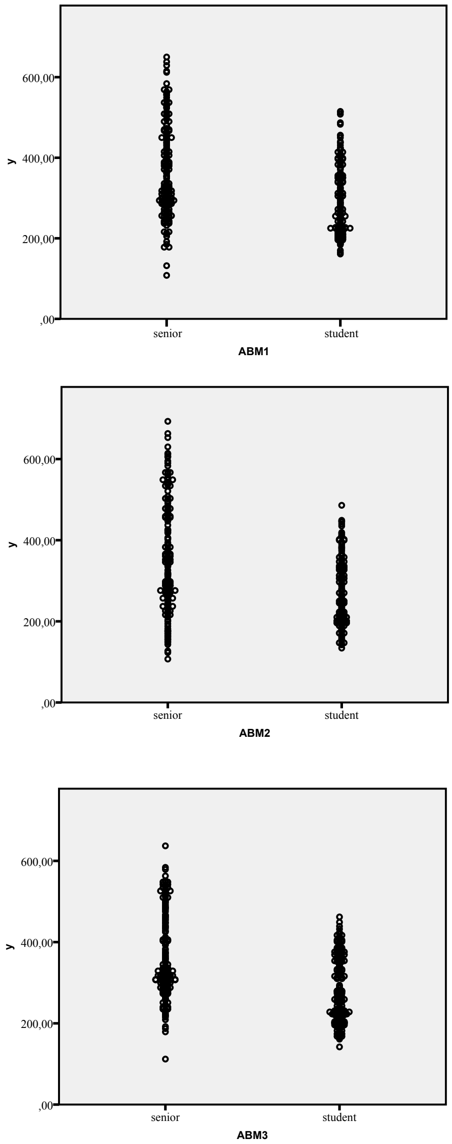

FIGURE 4.30: Figures of ABM 1-3 in virtual presence condition. Plots show that seniors have a tendency to look further down on the robot.

4.6 Eye-Tracker Summery

Cluster analysis 4.3 yielded that participants looked at the upper region of the robot, except for young VP ABM2, were no difference was found. Experiment 1 found that on all young in VP ABM and PP ABM1 participants fixated significantly higher on the robot compared to seniors, PP ABM2 showed no difference, with a small effect size, PP ABM3 the seniors fixated higher up on the robot, but also with a small effect size. On average the young participants fixated higher on the robot then the seniors4.5, this correlates with the findings of Nomura [20] and Wong [32], that people cross culturally fixate in the upper region of a humanoid figure.

4.7 Emotion Perception Results

In experiment 1 a total count of 54 participated over five test days; 34 female, 20 male. A total of 24 seniors participated; 16 female, 8 male. A total of 30 students participated; 18 female, 12 male. This and the following sections outline results of experiment one, first physical then virtual presence. Two male seniors are excluded from results due cognitive impairment. Two participants in the group of students are excluded from results for exceeding the upper age boundary.

4.8 Results for Physical Presence Condition

Participants: 14 students (5 male, 9 female) age range 17-29 (M=23 SD = 3.8 Mdn = 23). 13 seniors (3 male, 10 female) age range 67-89 (M = 74.6 SD = 6.5 Mdn = 75) participated in the physical presence condition. The tables 4.26, 4.27 and 4.28 shows the emotions each participant reported for each ABM andeach represents a robot movement where participants can choose multiple emotions. Table 4.26 shows, interestingly, that seniors are good at identifying anger compared to students. This does not reflect the findings of Nomura et al. [20]. The two-tailed probability of Fisher's exact test is reported.

[1] pugnacious
[2] relaxed, "mechanical"

ABM1 (anger)			
	Students (N = 14)	Seniors (N = 13)	Fisher's *p*
Anger	11 (78.6%)	9 (69.2%)	0.678
Hate	9 (64.3%)	6 (46.2%)	0.449
Happiness	4 (28.6%)	3 (23.1%)	1.000
Surprise	5 (35.7%)	6 (46.2%)	0.704
Sadness	4 (28.6%)	1 (7.7%)	0.326
Fear	8 (57.1%)	3 (23.1%)	0.120 (1-tailed p = 0.079)
Other	2[1] (14.3%)	2 (15.4%)	1.000

Table 4.26: Count of participants who identified given emotions after observing anger motion

ABM2 (happiness)			
	Students (N = 14)	Seniors (N = 13)	Fisher's *p*
Anger	1 (7.1%)	2 (15.4%)	0.596
Hate	1 (7.1%)	1 (7.7%)	1.000
Happiness	13 (92.9%)	9 (69.2%)	0.165
Surprise	10 (71.4%)	3 (23.1%)	0.021
Sadness	2 (14.3%)	1 (7.7%)	1.000
Fear	2 (14.3%)	3 (23.1%)	0.648
Other	4[2] (28.6%)	2 (15.4%)	0.648

Table 4.27: Count of participants who identified given emotions after observing happiness motion

ABM3 (sadnesss)			
	Students (N = 14)	Seniors (N = 13)	Fisher's *p*
Anger	2 (14.3%)	5 (38.5%)	0.209
Hate	2 (14.3%)	3 (23.1%)	0.648
Happiness	2 (14.3%)	3 (23.1%)	0.648
Surprise	5 (35.7%)	5 (38.5%)	1.000
Sadness	11 (78.6%)	4 (30.8%)	0.021
Fear	6 (42.9%)	5 (38.5%)	1.000
Other	5[3] (35.7%)	6[4] (46.2%)	0.704

Table 4.28: Count of participants who identified given emotions after observing sadness motion

Figure 4.31 shows the experimental setup at SOSU Nord where students participated.

Participants impression of speed and magnitude of the ABM 1, 2 and 3 were collected as Likert scale ratings of 1 (slow/small) to 5 (fast/large). Summary of motion speed is shown in table 4.29 while summary of motion magnitude is shown in table 4.30

[3]confused, "more empathic"
[4]sad, grumpy, hopelessness

FIGURE 4.31: Photo of 'physical presence' experimental setup at SOSU Nord where students were recruited. In front of the participant seat is the eye-tracker and the calibration board on the table.

FIGURE 4.32: Photo of 'physical presence' experimental setup at Vodskov where senior participants were recruited. On the table is the deactivated robot and the wall behind it has been cleared for paintings to avoid any visual points of interest. Remaining are only the nails.

Impression of speed		
	Students (N = 14)	Seniors (N = 13)
ABM1	M = 4.07. Std = 0.81	M = 3.23. Std = 0.72
ABM2	M = 3.57. Std = 0.66	M = 3. Std = 0.70
ABM3	M = 2.71. Std = 0.48	M = 2.76. Std = 0.59

TABLE 4.29: Summary of perceived motion speed (1-5 Likert rating) in the Physical Presence condition.

Impression of magnitude		
	Students (N = 14)	Seniors (N = 13)
ABM1	M = 3.42. Std = 1.03	M = 3.61. Std = 0.50
ABM2	M = 3.28. Std = 0.72	M = 2.92. Std = 0.49
ABM3	M = 2.64. Std = 0.76	M = 3.23. Std = 0.83

TABLE 4.30: Summary of perceived motion magnitude (1-5 Likert rating) in the Physical Presence condition.

4.8.1 Result of Emotion Recognition for Physical Presence

All the quantitative dependent variables summarized in table 4.1 are analyzed statistically. The perceived emotions are compared between young and seniors by looking for significant difference between how frequently a group perceives any of the seven options. This is done using Fisher's exact test and is performed for each of the three ABM expressions.

As expected the identification of happiness and sadness in table 4.27 and 4.28 is comparable with previous findings.

4.8.2 Result of Speed/Magnitude Perception for Physical Presence

Independent measures t-test are used to determine if any significant difference is present in the perception of speed and magnitude between students and seniors. Significant differences, and of good effect size, are found in speed perception of: ABM1; t(25) = -2.79, $p < 0.05$, r = 0.48. And ABM2; t(25) = -2.19, $p < 0.05$, r = 0.40. No significant difference are found for magnitude impressions. The graph 4.33 visualizes the contents of speed results in table 4.29. The graph 4.34 visualizes the contents of magnitude results in table 4.30. Calculations are located on the enclosed DVD in the folder 'Experiment 1\SPSS Output'.

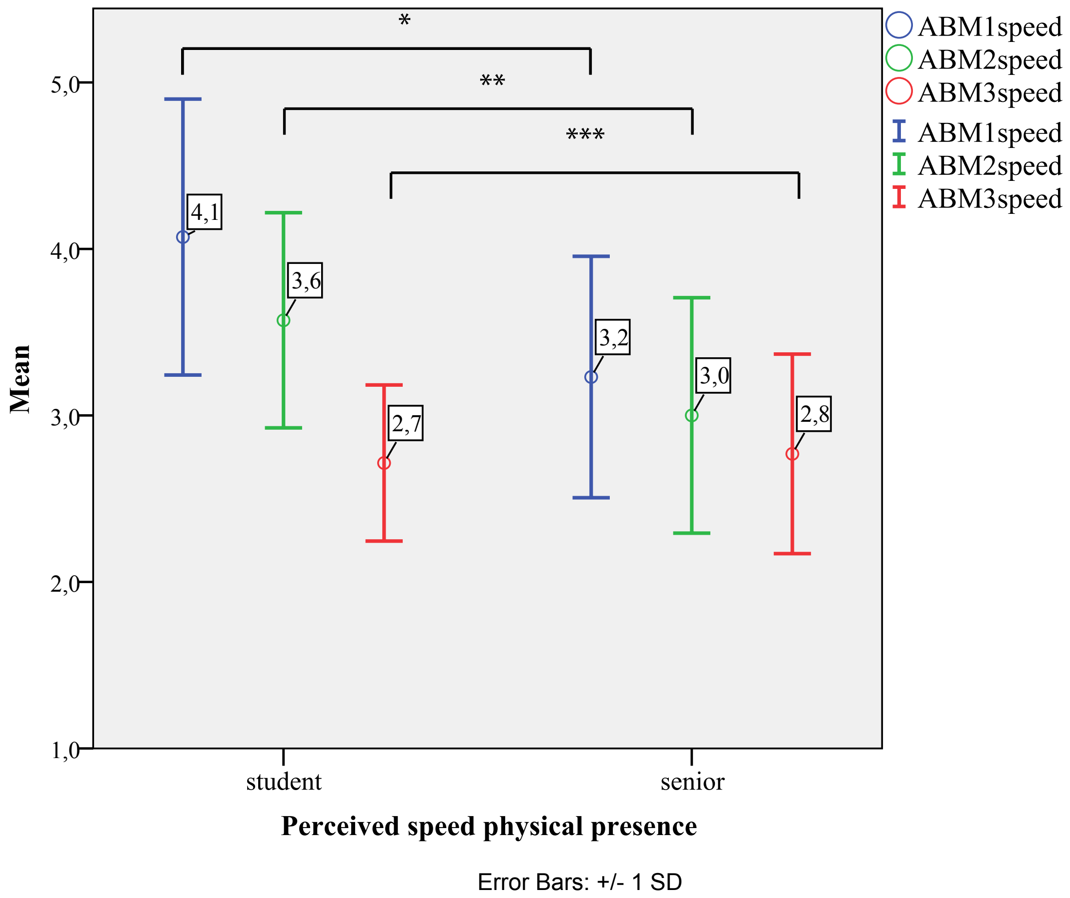

FIGURE 4.33: Representation of mean and standard deviation of how students and seniors perceived ABM speed. *P = 0.01, **P = 0.03, ***P >0.05

4.9 Results for Virtual Presence Condition

Participants: 16 students (7 male, 9 female) age range 17-31 (M=21,5 SD = 3,7 Mdn = 21,5). 11 seniors (5 male, 6 female) age range 64-77 (M = 70,3 SD = 3,7 Mdn = 70) participated in the virtual presence condition. The tables 4.31, 4.32 and 4.33 shows the set of emotions each participant reported for each ABM. Each table represents a video of one distinct ABM and participants can choose multiple emotions.

[5] enthusiasm
[6] agressive or happy
[7] nervous
[8] searching, withdrawn

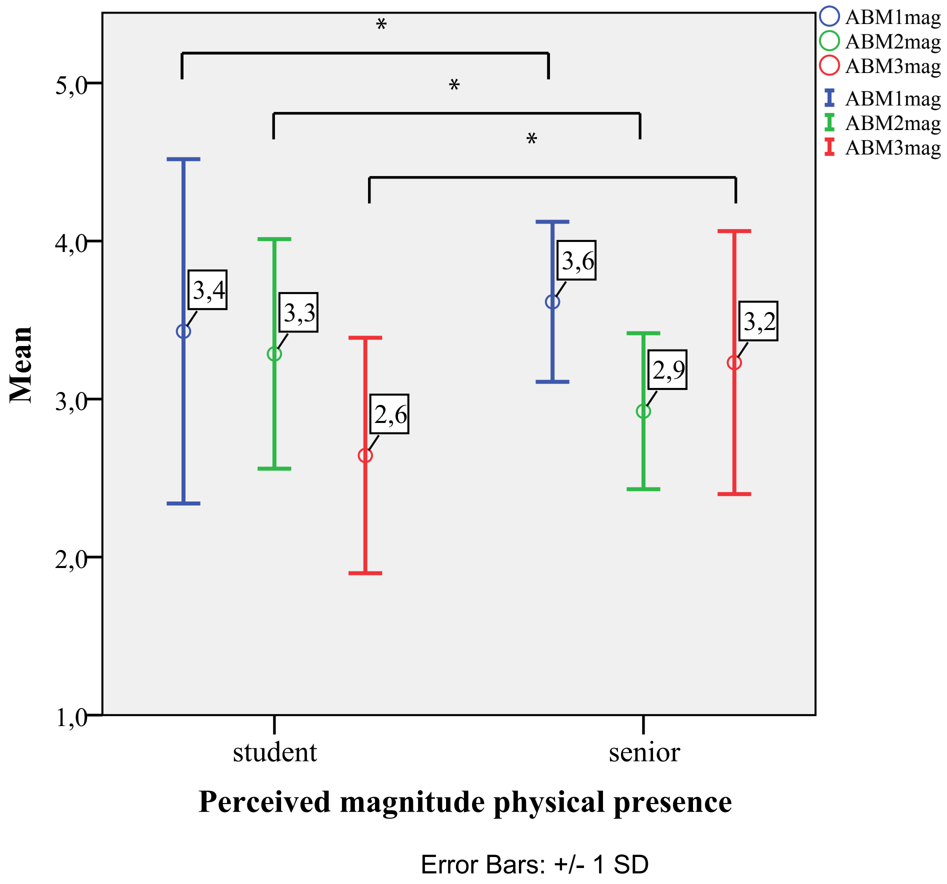

FIGURE 4.34: Representation of mean and standard deviation of how students and seniors perceived ABM magnitude. *P >0.05

FIGURE 4.35: Photo of 'virtual presence' experimental setup at SOSU Nord where students were recruited. Projector is located at a distance from the wall where the robot is shown in 1:1 scale.

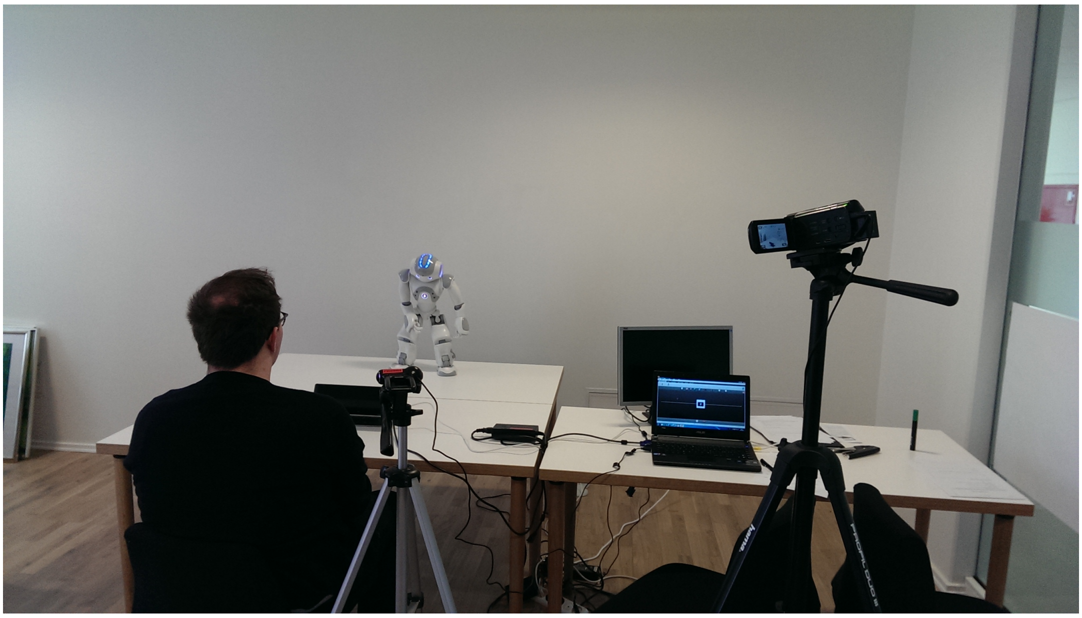

FIGURE 4.36: Photo of 'virtual presence' experimental setup at Fremtidens Plejehjem where seniors visited to participate. The robot was also used for two tests at this location.

ABM1 (anger)			
	Students (N = 16)	Seniors (N = 11)	Fisher's p
Anger	12 (75%)	7 (63.6%)	0.675
Hate	9 (56.3%)	3 (27.3%)	0.239
Happiness	6 (37.5%)	4 (36.4%)	1.000
Surprise	5 (31.3%)	5 (45.5%)	0.687
Sadness	3 (18.8%)	1 (9.1%)	0.624
Fear	2 (12.5%)	2 (18.2%)	1.000
Other	3[5] (18.8%)	2[6] (18.2%)	1.000

TABLE 4.31: Count of participants who identified given emotions after observing anger motion

ABM2 (happiness)			
	Students (N = 16)	Seniors (N = 11)	Fisher's p
Anger	0	3 (27.3%)	0.056
Hate	0	2 (18.2%)	0.157
Happiness	13 (81.3%)	6 (54.5%)	0.206
Surprise	8 (50%)	7 (63.6%)	0.696
Sadness	3 (18.8%)	3 (27.3%)	0.662
Fear	3 (18.8%)	4 (36.4%)	0.391
Other	1 (6.3%)	2 (18.2%)	0.549

TABLE 4.32: Count of participants who identified given emotions after observing happiness motion

	ABM3 (sadnesss)		
	Students (N = 16)	Seniors (N = 11)	Fisher's p
Anger	2 (12.5%)	2 (18.2%)	1.000
Hate	0	2 (18.2%)	0.157
Happiness	2 (12.5%)	3 (27.3%)	0.370
Surprise	2 (12.5%)	6 (54.5%)	0.033
Sadness	14 (87.5%)	2 (18.2%)	0.001
Fear	7 (43.8%)	1 (9.1%)	0.090
Other	1[7] (6.3%)	4[8] (36.4%)	0.125

Table 4.33: Count of participants who identified given emotions after observing sadness motion

4.10 Result of Speed/Magnitude Perception for Virtual Presence

A significant difference in speed impression is present between students and seniors in ABM3 with good effect size; t(25) = -2.87, $p < 0.05$, r = 0.49. In the magnitude impression a significant difference is present between students and seniors in ABM3 with good effect size; t(25) = -2.57, $p < 0.05$, r = 0.45. The graph 4.37 visualizes the contents of speed results in table 4.34. The graph 4.38 visualizes the contents of magnitude results in table 4.35.

Participants impression of speed and magnitude of the virtually present ABM 1, 2 and 3 were reported as Likert scale ratings of 1 (slow/small) to 5 (fast/large). Summary of motion speed is shown in table 4.34 while summary of motion magnitude is shown in table 4.35

	Impression of speed	
	Students (N = 16)	Seniors (N = 11)
ABM1	M = 3.75. Std = 0.77	M = 3.54. Std = 0.52
ABM2	M = 3.18. Std = 1.04	M = 3. Std = 0.44
ABM3	M = 2.25. Std = 0.85	M = 3.09. Std = 0.53

Table 4.34: Summary of perceived motion speed (1-5 Likert rating) in the Virtual Presence condition.

Calculations are located on the enclosed DVD in the folder 'Experiment 1\SPSS Output'.

Impression of magnitude		
	Students (N = 16)	Seniors (N = 11)
ABM1	M = 3.06. Std = 0.92	M = 3.36. Std = 0.50
ABM2	M = 2.8. Std = 0.75	M = 3. Std = 0.63
ABM3	M = 2.43. Std = 0.81	M = 3.18. Std = 0.60

TABLE 4.35: Summary of perceived motion magnitude (1-5 Likert rating) in the Virtual Presence condition.

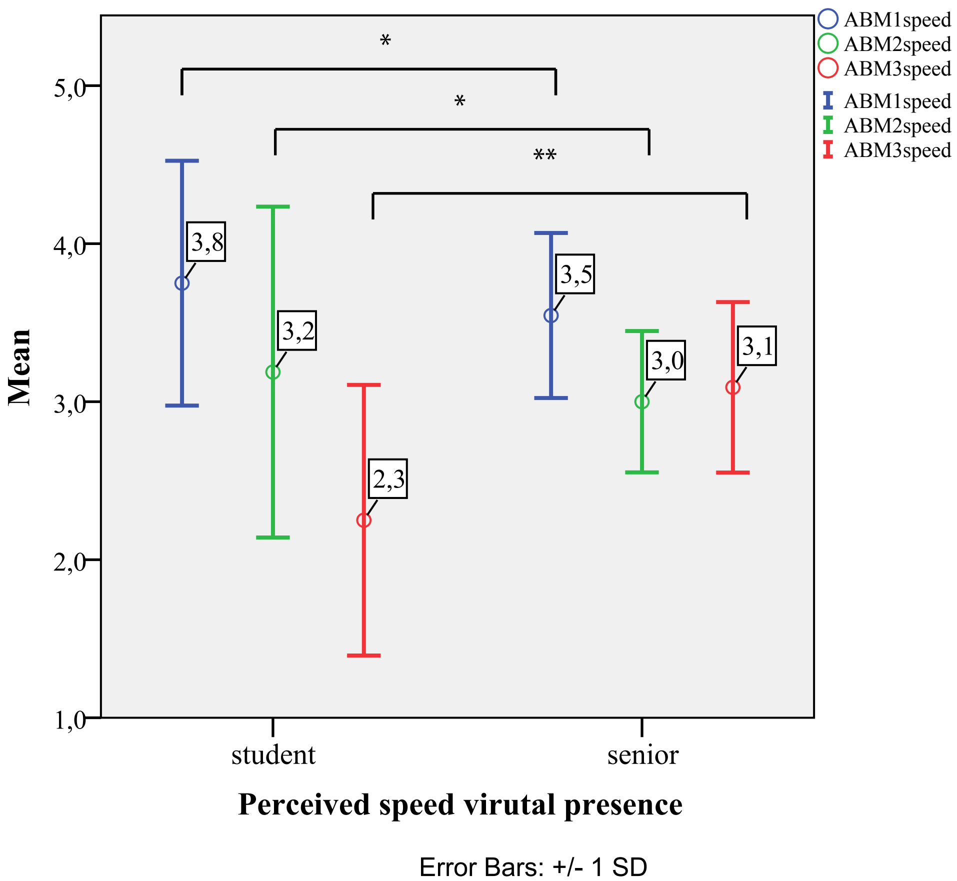

FIGURE 4.37: Representation of mean and standard deviation of how students and seniors perceived ABM speed. *P >0.05, **P = 0.008.

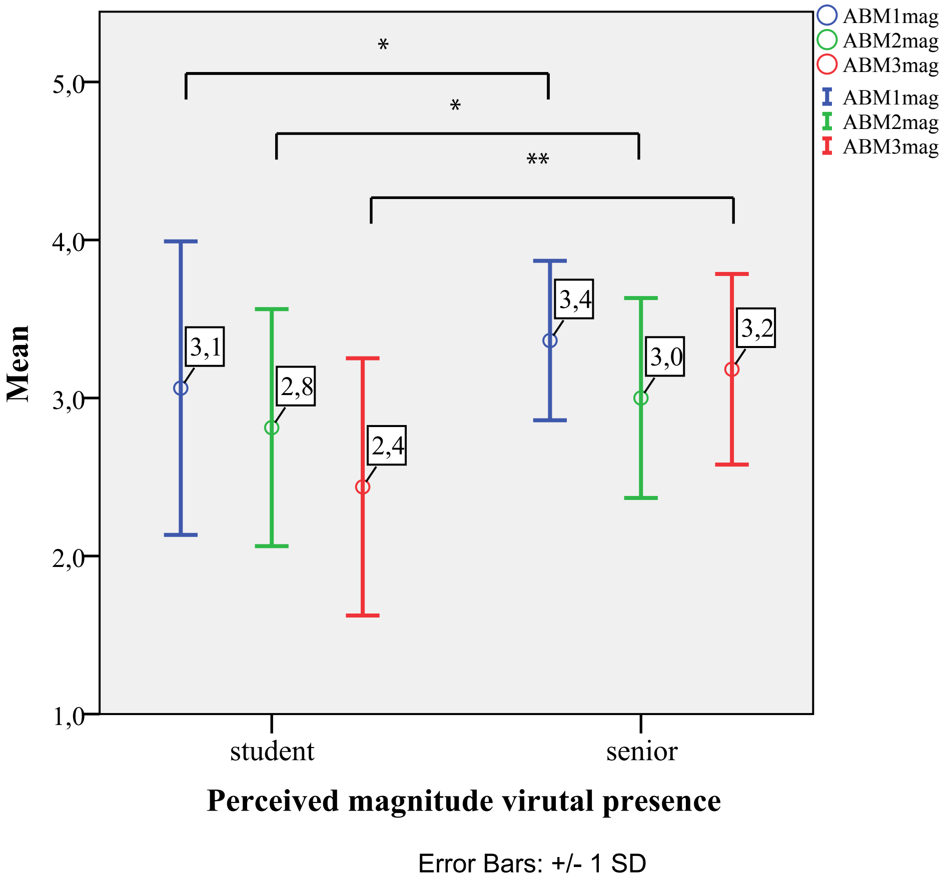

FIGURE 4.38: Representation of mean and standard deviation of how students and seniors perceived ABM magnitude. *P >0.05, **P = 0.016.

4.11 Analysis

This section presents a summary of the results from test 1 and goes on to analyse them. In relation to research question 1 table 4.36 gives a direct comparison of results from emotion recognition between [20] (table 4 in the Nomura paper) and section 4.8 of this report; tables 4.26, 4.27 and 4.28. The results for happiness and sadness shows that Danish and Japanese seniors are comparable in their perception of happiness and sadness in body motion, but differ in the recognition of anger. Here Danish participants are not significantly different from their younger counterparts.

Emotion Identification Comparison			
	Nomura [20]	Experiment 1	Compatibility
Anger	Significant ($p = 0.000$)	Insignificant ($p = 0.678$	No
Happiness	Insignificant ($p = 0.319$)	Insignificant ($p = 0.165$)	Yes
Sadness	Significant ($p = 0.038$)	Significant ($p = 0.021$)	Yes

Table 4.36: Difference in performance between students and seniors. Summary of perceived emotions compared to the results of Nomura. In anger (ABM1) Danish seniors were on level with Danish students at identifying the emotion.

In relation to research question 2. table 4.37 show that the presence of the robot did not significantly alter emotion perception of the intended emotions. As in the co-location condition: No significant differences are found in anger and happiness but a difference exists in sadness recognition.

Emotion Identification Comparison			
	Experiment 1 (virtual)	Experiment 1 (physical)	Compatibility
Anger	Insignificant ($p = 0.675$)	Insignificant ($p = 0.678$	Yes
Happiness	Insignificant ($p = 0.206$)	Insignificant ($p = 0.165$)	Yes
Sadness	Significant ($p = 0.001$)	Significant ($p = 0.021$)	Yes

Table 4.37: Difference in performance between students and seniors. Summary of perceived emotions compared within experiment 1 between physical and virtual presence. The relative ability of student and seniors to identify emotions in Nao are comparable.

Hypothesis 1: "Danish seniors will have lower accuracy of emotion perception than young" is retained. The difference in gaze between seniors and students which was found by analyzing distributed on the y-axis shown in chapter 4.5 correlates with the sources in [32]: An age-dependent deviation exists in gaze patterns. In [20] the performance of anger perception for Japanese seniors was found to be lower than students. Surprisingly this was not the case for Danish seniors who were on level with Danish students. It is possible that a cultural difference exists. Alternatively the ABM design from section 3 is easily recognized by seniors. In the second part of experiment 1 the change in presence from a co-located robot to a virtual robot did not change emotion perception. The reason is likely that participants passively observe the robot (instead of talking to it, e.g.) and nothing obscures their view. Hypothesis 2: "The presence of the robot affects emotion recognition accuracy of both groups of participants." is retained. The immediate results for emotion recognition in the virtual presence condition are comparable to those of physical presence. Seniors are not significantly less capable of identifying anger and happiness. A significance is present in perception of sadness where students identified the emotion more often than seniors.

Chapter 5

Contextualized ABM - The ”Robot Game”

The context of a game simulates intrinsic motivation in the robot to win, thus emotional reactions become justifiable. This is also discussed in section 1.2. A proof-of-concept behavior is developed for the robot to enable it to play a game with the user. It is fully automated so the final test is not Wizard of Oz. The game is played with dice where Nao will guess the outcome of each subsequent roll and estimate if the next roll will be above or below the previous one. The robot wins the game if it can guess correctly 2 out of 3 times. Behavior of the robot is programmed in Choregraphe and is performed dynamically in response to game event. Computer vision (CV) is used to recognize the outcome of the dice roll. Due to limitations in memory and processing power of the of the robot CV is performed separately on a laptop.

5.1 Programming the Robot

This section describes the methods used to program a system for the NAO robot to play a dice game. Python based programs are implemented on a PC in order to perform CV for the robot and return the analysis results back to the it in real-time. To play the game the following hardware is necessary: A robot, a Windows PC, a white/black six-sided dice, a green surface of A4-paper dimensions placed on a static surface with a height of 16cm. Source code is located in the enclosed DVD in folder ’Robot game’. Consult

README file for more details. Figure 5.1 shows the basic setup. The AV production on the enclosed DVD can be seen as supplement the descriptions of the robot found in this chapter.

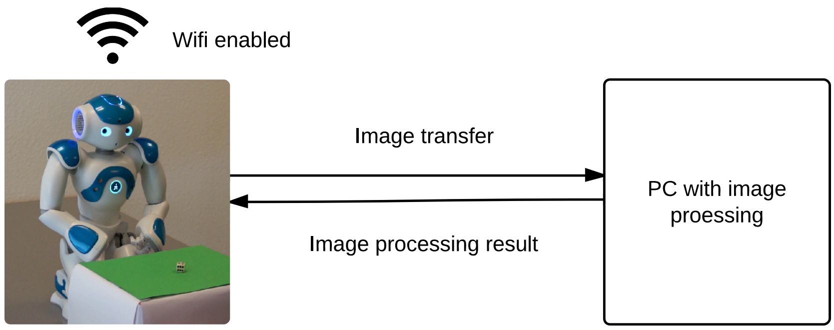

FIGURE 5.1: Relation between robot and PC to perform computer vision tasks.

System details: Robot: Nao H25, head v3.3 and v4 running OpenNAO (Gentoo based) OS version 1.14.5. Laptop: 2.3GHz i7 x64 CPU, 8GB ram running Windows 7 , 64 bit. Software: Choregraphe 1.14.5 for programming the robot graphically. Python 27 is used within Choregraphe for greater control. SimpleCV 1.3 library for image processing in Python. A TCP socket connection programmed in Python is used to transfer images and dice analysis result between PC and the robot and all Python programming outside Choregraphe is done the Python IDE; IDLE. For recording voice a ZOOM H4n digital voice recorder is used. For modifying the voice recordings Audacity v2.0.3 is used.

In its current version the robot does not support Danish speech. Therefore voice is recorded and modified to sound "robotic" and fit the standard synthetic voice of the Nao robot. Ad-hoc modification is performed on recorded voice with the following parameters: pitch shifted up by 7 semitones. Tremolo effect at 10hz + "wet level" 17%.

5.1.1 Computer Vision

Choregraphe supports vision recognition by learning, but upon trial an ordinary white six-sided dice with back dots does not give enough variety to be stored in the recognition database on the one/two-dot sides of the dice. The robot is programmed to position its camera (head) over the dice and capture a top-down image for analysis. This lowers that complexity of analysis as there is almost no other shapes present in an image that must be taken into account when identifying the roll of the dice. The dice recognition task is done by programming an image analysis application using SimpleCV in Python and the primary resource for using the library is [7] and online documentation from the developer [27]. This program is run on Windows 7 and images are send from the robot to a PC, analyzed and the result is send back to the robot. Various approaches are tested to perform segmentation on images which varies in performance based on the light conditions and distance from camera to dice. The chosen method is to binarize the input image based on the green color channel. A function of the SimpleCV library is used for blob detection[1] and the exact parameters can be found in the source code. A change from robot head version 3.3 to 4 required changes in the code as the heads use different camera versions. The newer taking pictures with higher field of view and color saturation. The original image processing method use a more complicated process of image optimization before blob detection. After cropping the area of interest the following functions are used: white balance, dilation, erosion, morphology-close and finally blob detection. Finally the array of blobs is accessed and blobs of correct size is counted and the count is returned. Images taken with the newer head can be processed with very few steps as presented in figure 5.2, and in fact the initial process caused so much distortion to images that the final blob count was too unreliable.

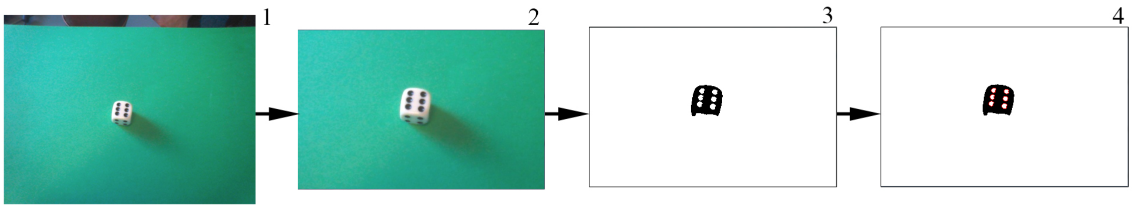

FIGURE 5.2: Stages of image processing for dice analysis. 1: Original image. 2: Crop. 3: Binarize using green filter. 4: Blob detection with size threshold (red circles indicate detected blobs).

A second computer vision function is designed but not completed. The "findBoard" script allows the robot to estimate its location relative to the position of the green game board. This functionality would allow the robot to walk from its position in front of the board, perform an ABM and relocate itself in front of the board automatically. In the tested version of the game system slight correction of the robots sitting position is needed after an ABM is performed. Alternatively the "findBoard" script could be used to position the head directly over a dice on the green board rather than relying on a repetitive animation and demand for the dice to be rolled in the center of the

[1]The process of locating binary areas of interest

board for each roll analysis. The script relies on edge detection to find the green board, find the backmost corners of the board and the calculate the distance from 2D to 3D. Knowing the distance between two points in 2D and the focal length of the head camera the Pythagorean theorem can be used to calculate the board-camera distance. It does require the head camera to be at the perpendicular center point between the two points which requires the robot to positions its body and head by walking, based on inclination of the board lines found through edge detection. Mathematical calculations to solve this task are included on the DVD in the 'Robot game' folder as image files "findBoard calculations" A/B.

5.1.2 Socket Connection

The standard Python library 'socket' is used and is structured as server (PC)/client (robot). The server program must be running on a PC first. When the Choregraphe program on the robot needs to read the value of a dice roll a Python script box in Choregraphe takes a picture, connects to the the PC by its IP address and then sends a command string. Command strings are implemented to make the code modular and support several image processing features but in the tested version of the system only one command is accepted; "findBlob". The server sends back the string reply "sendImage" to indicate the server is online and ready. A loop function "reciveImage" receives the data stream and writes the image file to the C:/ drive until the transfer is complete. The server continues to call the image processing script from a separate file which returns an integer with the number of detected dice eyes. Finally the server sends this integer back the Choregraphe program on the robot and ends the connection.

5.1.3 Game Rules and Behavior

A dice game is programmed in Choregraphe for participants to play with the robot. The objective is for the robot to guess if the outcome of a dice roll is higher or lower than the previous roll. The participants involvement is to roll the dice for the robot and trigger it to look at the dice through hand-sensor touch. Each game consists of three rounds with each round containing four throws of the dice where the robot makes three guesses. For the robot to win a round it must guess correctly at least twice. To win the game the robot must win at least two rounds. If the robot wins a round it displays happiness (ABM 2) and anger (ABM1) if it loses a round. If the robot wins the game it displays happiness and sadness (ABM3) if it loses the game. Some basic functionality of the game rules are outlined in figure 5.3

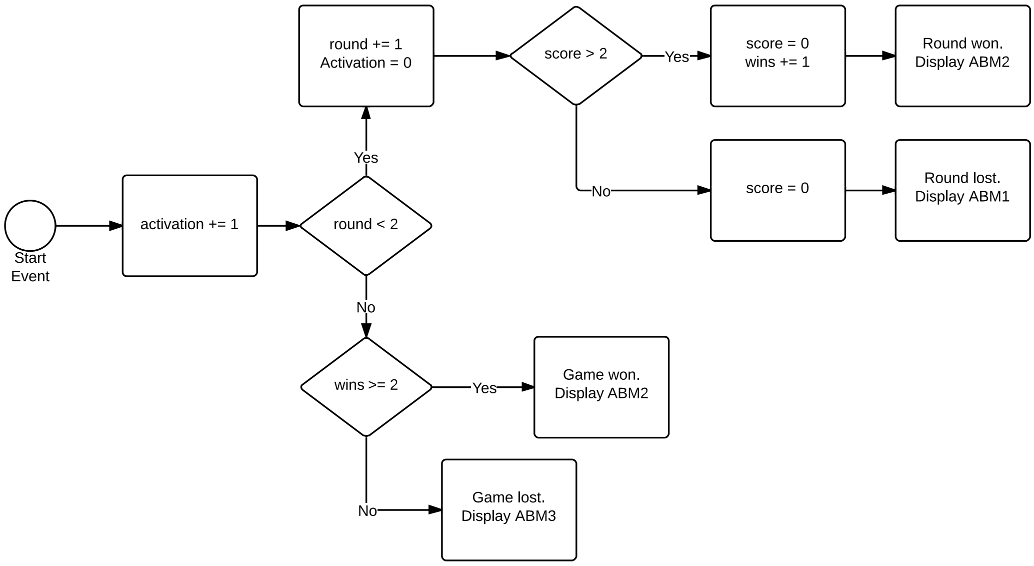

FIGURE 5.3: Flowchart is an except of the game rules. Code executed every time the robot has taken a picture. Variables incremented or set to zero are: Activation; increments every time robot is triggered. It keeps count of turn number. Round; number of rounds played. Score; incremented in a separate function. Wins; counts the number of won rounds.

A simple survey is conducted to decide the behavior of the robot. Four people are observed using camera recordings while they play the game with the observation conductor. Special attention is taken to body language of the player. All players were familiar with the "dealer". Noticeable behavior was their gaze direction which was focused on the dealer when making the higher/lower guess. Gaze was directed at the table towards the dice most of the time. Minor movement of head and arms and small vocalizations in between guessing. The gaze of the player was also drawn by movement of the dealer and their level of interaction was minimal. For the most engaging experience of playing with the robot all of these behavioral traits should be implemented but due to time constraints the following were not implemented in the robot: subtle vocalizations, subtle movement of the body while waiting, gaze drawn by movement of player.

Chapter 6

Experiment 2

To investigate how contextualization affects emotion perception this experiment asks participants to play a game with the robot and observe how the robot reacts to winning and losing.

6.1 Design

The experiment is structured the same way as experiment 1. Independent variables are the same ABM as in experiment 1. Dependent variables are the perception of ABM and magnitude/speed. The questionnaire used to measure emotion, magnitude and speed are reused from experiment 1. A group of seniors are asked to play a dice game with the robot and the robot will then give an ABM reaction that relates to its luck. The participants has no stake in the game and cannot win or lose. Participants are taught the rules of the game, as described in section 5.1.3 and asked to roll the dice for the robot. The robot performs an ABM after each round and the participant must then fill out the perception questionnaire (see Appendix A) before the game continues. See figure 6.1 for setup description and overview. Because the outcome of the game dictates what ABM will be displayed the same ABM may be shown several times. In this case the participant is not given a questionnaire to rate the same ABM again. The hypothesis for experiment 2 is:

Hypothesis 3. *Danish seniors will have a higher accuracy of emotion recognition in the context of a game than if no context is given.*

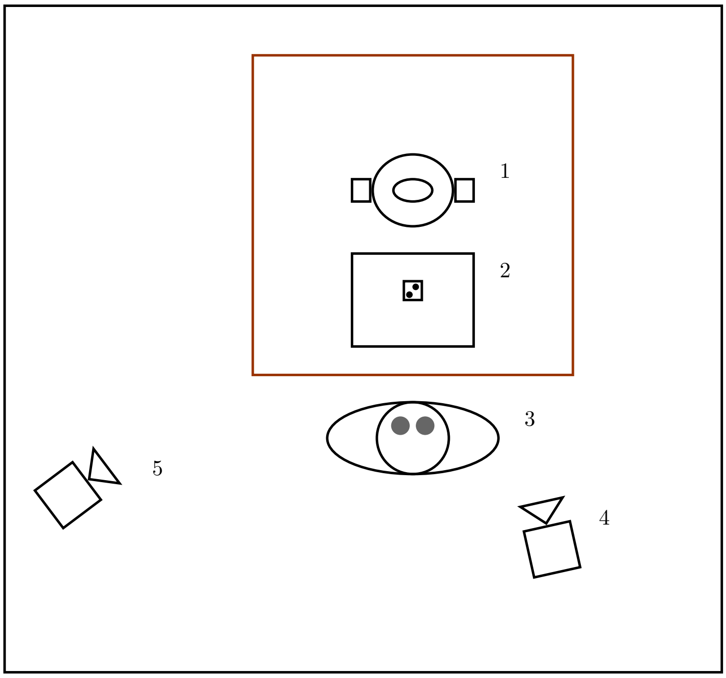

Figure 6.1: Illustration of the experiment setup seen from a top-down perspective. Scale not one-to-one. 1: Robot on table. 2: Gaming board and die on table. 3: Participant. 4: Video camera 1. for experiment recording). 5: Video camera 2. for experiment recording.

6.1.1 Participants

Senior citizens who either attend activities at the center or are otherwise from the area, are recruited by employees. A screening process is not used but employees are acquainted with individuals who they recruit. The same requirements as experiment 1 are used: age 64 or above and ”mentally fit” e.g. free of impairing cognitive affliction such as dementia. Additionally participants cannot participate in experiment 2 if they participated in experiment 1.

6.1.2 Apparatus

Nao H25 (v4 head) robot by Alderbran Robotics was used. Its colors are white with metallic blue plates, unlike the grey plated robot used in experiment 1. A Canon Legria HFM506 camera for recording the experiment. Tripod for stabilizing the camera. Laptop for image processing, specification described in section 5.1, Wifi router for robot-laptop connectivity.

6.1.3 Experiment 2 Procedure

Experiment 2 can be described in the same segments as experiment 1: introduction, test, interview. Participants are greeted to the test and given a short presentation of the project and the condition of their participation. They are asked to sign a consent form and encouraged to ask questions whenever such may occur. After the introduction participants are explained that they will be handed a questionnaire after the robot performs movement while standing up. Before the test starts participants are taught the rules of the game by playing it with the test conductor in place of the robot. Participants are also instructed how to trigger the robot between each throw and are explained this step is necessary because the robot does not "see very well". When the game is underway participants are instructed to fill out the questionnaire for emotion perception and movement speed/magnitude impression. After playing three rounds of the game a small interview is conducted. Finally the participant is thanked for participating and greeted farewell and next participant is welcomed.

6.1.4 Experiment 2 Pilot Test

A pilot test of experiment 2 is conducted, to evaluate the feasibility, the time of the experiment and to improve upon the experiment design and robot behavior. The pilot test used the same procedures described in section 6.1.3. The pilot test involved three 10th semester medialogy students and two 4th semester medialogy students. The five students who were involved in the study, confirmed the test and that test procedure worked. However the importance of turning up the volume of the robots voice and adding feedback to the user during each turn, was mentioned by the participants, and noted. One also suggested that the player should have pen and paper to keep track of the score.

6.2 Experiment 2 Results

12 seniors (8 female and 4 male) participated in experiment 2. Two participants were 63 years old but were not rejected from testing since the age limit of 64 is arbitrary and the number of participants is low. The exact age of two female participants is unknown.

Age ranged from 63-81 years (M = 70.7, SD = 5.9, Mdn = 70). The robot lost more rounds than it won and as a result only half the participants saw a display of happiness. Figure 6.2 shows a photo of the experiment setup. Therefore the number of participants in table 6.2 is N = 6 while N = 12 in tables for anger 6.1 and sadness 6.3. Results of perceived emotions are compared to the results of seniors in the physical presence condition of experiment 1. The 2-sided significance of Fisher's exact test is reported.

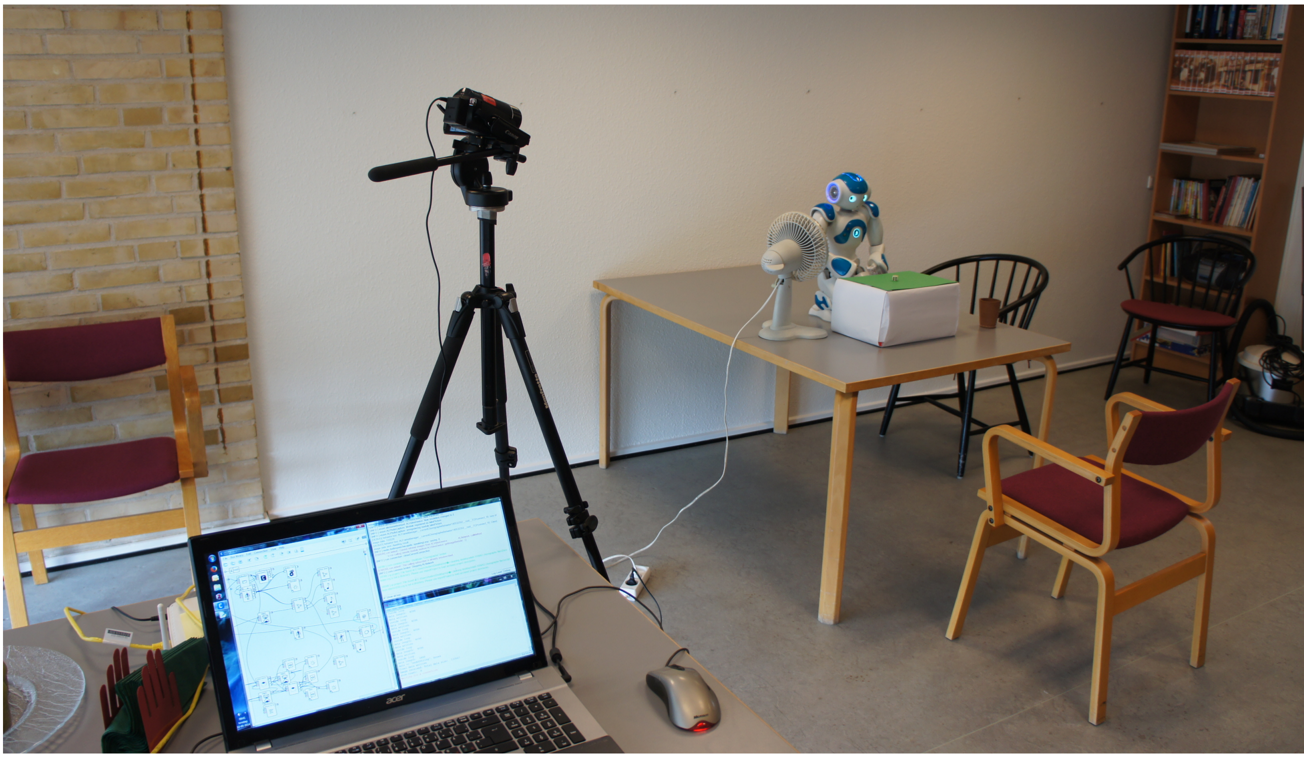

FIGURE 6.2: Photograph of the test setup at Vodskov with the robot on a table, ready tpp play. The fan shown in the picture was removed during tests.

ABM1 (anger)			
	Contextualized (N = 12)	Decontextualized (N = 13)	Fisher's p
	Count	Count	
Anger	6 (50%)	9 (69.2%)	0.428
Hate	4 (33.3%)	6 (46.2%)	0.688
Happiness	5 (41.7%)	3 (23.1%)	0.411
Surprise	8 (66.7%)	6 (46.2%)	0.428
Sadness	1 (8.3%)	1 (7.7%)	1.000
Fear	1 (8.3%)	3 (23.1%)	0.593
Other	8[1] (66.7%)	2 (15.4%)	0.015

TABLE 6.1: Count of participants who identified given emotions after observing anger motion. Compaison is between participants in experiment 2 and experiment 1 (decontextualized); seniors, physical presence.

[1] "betænkelighed"

ABM2 (happiness)			
	Contextualized (N = 6) Count	Decontextualized (N = 13) Count	Fisher's p
Anger	2 (33.3%)	2 (15.4%)	0.557
Hate	1 (16.7%)	1 (7.7%)	1.000
Happiness	4 (66.7%)	9 (69.2%)	1.000
Surprise	3 (50.0%)	3 (23.1%)	0.320
Sadness	0 (0.0%)	1 (7.7%)	1.000
Fear	2 (33.3%)	3 (23.1%)	1.000
Other	2 (33.3%)	2 (15.4%)	0.557

Table 6.2: Count of participants who identified given emotions after observing happiness motion

ABM3 (sadnesss)			
	Contextualized (N = 12) Count	Seniors (N = 13) Count	Fisher's p
Anger	4 (33.3%)	5 (38.5%)	1.000
Hate	2 (16.7%)	3 (23.1%)	1.000
Happiness	7 (58.3%)	3 (23.1%)	0.111
Surprise	8 (66.7%)	5 (38.5%)	0.238
Sadness	2 (16.7%)	4 (30.8%)	0.645
Fear	3 (25.0%)	5 (38.5%)	1.000
Other	5[2] (35.7%)	6[3] (46.2%)	0.704

Table 6.3: Count of participants who identified given emotions after observing sadness motion

Impressions of motion speed and magnitude for the contextualized ABM there was no significant difference compared to the decontextualized counterpart. Independent t-test is used to compare means of speed and magnitude ratings of each ABM. All calculations are located on the enclosed DVD.

6.3 Analysis

By comparison with seniors under the same condition of physical presence, contextualization did not improve emotion recognition. A lower number of seniors identified each ABM as the emotion they are designed to express. While the results of experiment 2 suggests that accuracy of emotion identification becomes lower, there might be explanations to this: Even though the game is simple it is not easy to keep account of the

[2] gymnastic movement, snooty
[3] sad, grumpy, hopelessness

point system. Therefore the senior participants probably did not know if the robot was winning or losing. This problem became apparent during the pilot study and to solve it vocalizations were added to the robot after each round. But it is assumed that all participants were unable to hear the vocalization due to low volume and unclear pronunciation. In addition to confusion about the state of the game the ABMs are not shown in random order and as a result sequential bias is a strong possibility. Anger is most frequently the first ABM to be displayed as the robot often lost rounds. Based on feedback from participants during tests it can be assumed that the first ABM they saw (all tests, all conditions, all participants) was rated with higher inaccuracy. This relates to the simple matter of practice and improving at solving a task. This is a possible explanation to why 8 reported "other" emotion in ABM1. Hypothesis 3 for experiment 2; "Danish seniors will have a higher accuracy of emotion recognition in the context of a game than if no context is given." is rejected.

Chapter 7

Dementia Participant

During experiment 1's test sessions a senior person diagnosed with vascular dementia volunteered to participate in Experiment 1 see image 7.1, because of the diagnoses the participant was not included in experiment 1's main group. Diagnosed with Vascular Dementia (VD) the participant could yield hints, in how a person suffering from dementia symptoms perceive expressed ABM emotions. However it is important to note that since only one VD participant participated, and that the main purpose of the project was not to test on participants that was diagnosed with VD, no conclusions was drawn from such a small sample size. The VD participants followed the same procedure described in section 4.1.3 and the same statistical procedures described in section 4.3.3 were applied, with the exception of applying the KM2 procedure on the data set, because of the small sample size.

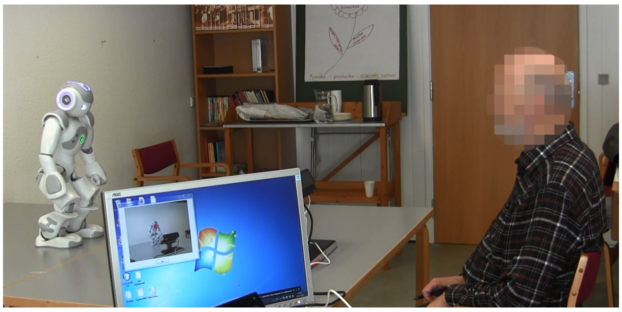

FIGURE 7.1: Photo of VD participant watching ABM1, face blurred out to hide identity.

VD Participant Answers to Questionnaire

The VD participant in addition answered the questionnaire perceived ABM emotions.

VD Participant ABM1

The VD participant answered "I do not think so" for all the emotions in ABM 1 (anger), except for "happiness" which he answered "I am certain" to that he saw. For the magnitude and speed the answered 3 on the liket scale for both.

VD Participant ABM2

The VD participant answered "I dont think so" for all the emotions in ABM2 (happiness), execpt for "happiness" which he answered "I think so" to that he saw. For the speed he answered 4 and for the magnitude he answered 3, on the likeret scales.

VD Participant ABM3

The VD participant answered "I do not think so" for all the emotions in ABM3 (sadness), except for "other" which he was "I am certain " that he saw, however the participant was not able to state precisely what he saw, do to cognitive failure. For ABM 3 the participant responded the magnitude and speed as 3 on the likeret scale both.

7.1 VD ABM discussion

The results gathered from the VD participant perceived ABM's is too small to analyse. However it was interesting to see the VD participant answered that he strongly perceived ABM1 as happiness, however he also answered that he thought he saw happiness in ABM2 . I could be that the two movements have simulates in that the robot moves more around in ABM1 and 2, then in ABM3. however ABM1 and 3 are misidentified and could be in line with [32] , Of cause with only one participant and a such could also be coincidences. The small sample size of the perceived magnitude and speed of the emotions are not calculated, do their size.

7.2 VD Eye Tracker-Results

This section analyse and discuss the eye-tracker results from VD participant.

ABM 1 Senior Dementia Participant

KM1 yielded one cluster containing 11 fixation marks (cluster 1) with the x,y-means (310.5, 185.3) and another cluster containing 23 fixation marks (cluster 2) with the x,y-means (296.5, 196.2), presented in table 7.1 and plotted in 7.2. The chi-square test yielded a significant difference with df $\bar{1}$ and a p-value $\bar{0}.03959$, between the upper or lower region of where the VD participant looked at the robot. The follow up binomial test gave a significant result with p$\bar{0}.05761$ with a probability of success 67 % that the participant looked at the lower part of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	310.5	185.3	11
Cluster 2	296.5	196.2	23

TABLE 7.1: KM1 x,y-means and number of fixation marks from two clusters from senior VD participants-in physical presence ABM1 - observing anger motion

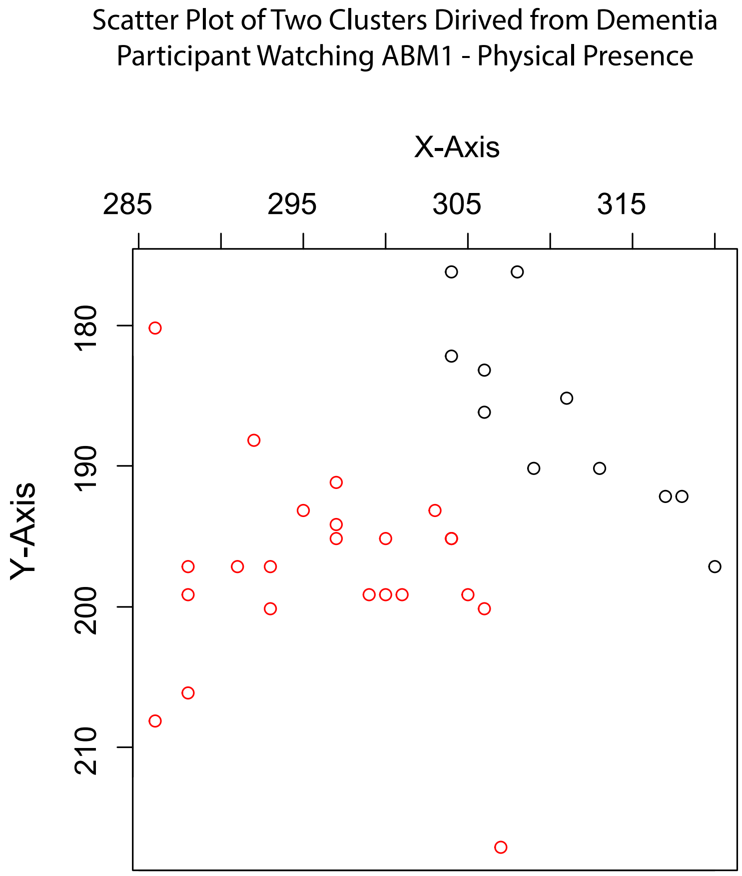

FIGURE 7.2: 2D data plot of two clusters from VD senior participant in physical presence ABM1 - observing anger.

ABM2 Senior Dementia Participant

KM1 yielded one cluster containing 2 fixation marks (cluster 1) with the x,y-means (330.0, 227.5) and another cluster containing 16 fixation marks (cluster 2) with the x,y-means (309.7, 178.5), presented in table 7.2 and plotted in 7.3. The chi-square test yielded a significant difference with df= 1 and a p-value = 0.0009674. The follow-up binominal test yielded a significant p-value = 0.001312 with a probability of success 88%, that the VD participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	330.0	227.5	2
Cluster 2	309.7	178.5	16

TABLE 7.2: KM1 x,y-means and number of fixation marks from two clusters from senior VD participants-in physical presence ABM2 - observing happiness motion

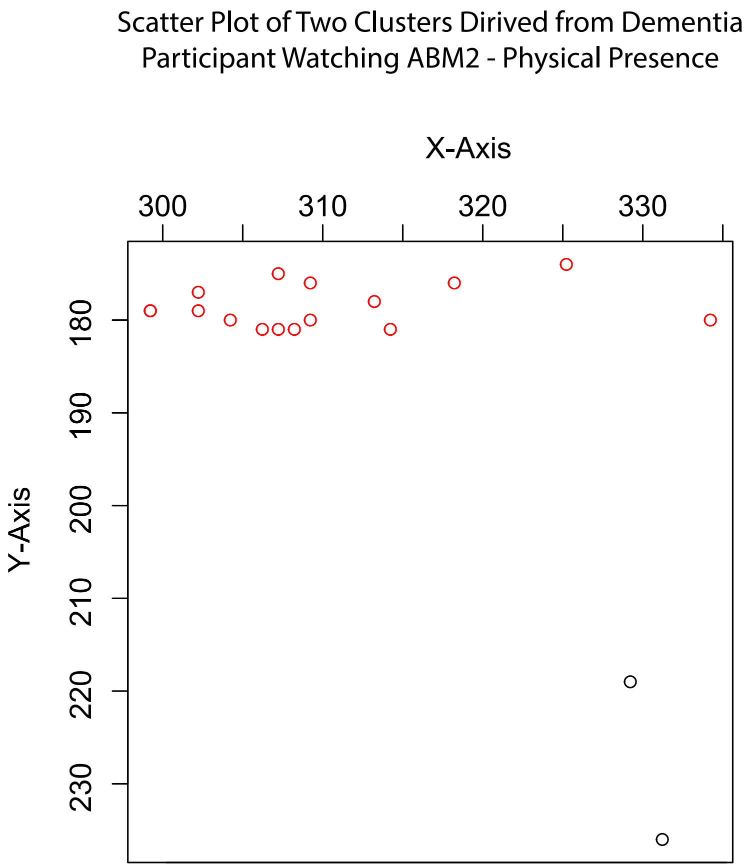

FIGURE 7.3: 2D data plot of two clusters from VD senior participant in physical presence ABM2 - observing happiness.

ABM3 Senior Dementia Participant

KM1 yielded one cluster containing 29 fixation marks (cluster 1) with the x,y-means (280.7, 193.3) and another cluster containing 2 fixation marks (cluster 2) with the x,y-means (293.0, 285.5), presented in table 7.3 and plotted in 7.4. The chi-square test yielded a significant difference with df = 1 and a p-value = 1.239e-06. The follow-up binominal test yielded a significant p-value = 4.629e-07 with a probability of success 93%, that the VD participant looked at the upper region of the robot.

Clusters	X-coordinate mean	Y-coordinate mean	number of fixation points
Cluster 1	280.7	193.3	29
Cluster 2	293.0	285.5	2

TABLE 7.3: KM1 x,y-means and number of fixation marks from two clusters from senior VD participants-in physical presence ABM2 - observing sadness motion

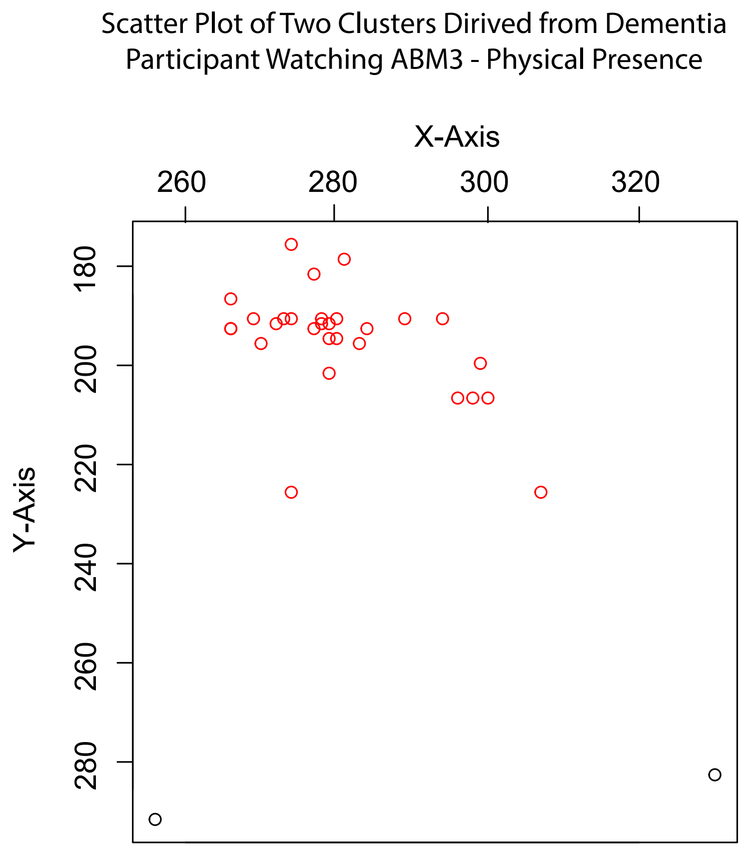

FIGURE 7.4: 2D data plot of two clusters from VD senior participant in physical presence ABM3 - observing sadness.

7.3 Discussion of VD Eye-Tracker Results

The small sample size makes it impossible to make any conclusions on where VD participants looks at robots. However comparing the VD participant with the mentally fit senior participants from ABM1,2,3, the fixations plots of the VD participant generally seems to focus at the upper region of the robot when looking at the fixation marks coordinates in general, this is consistent with 2.1.3 description that VD participants rely more on body language than normally mental fit senior participants. As such even though there a statistical difference between where the VD participant looks at either the upper or lower region of the participant.

Chapter 8

Experiment Interviews

This section centres on the answers gathered from the interviews in experiment 1 and 2. The participants were asked two opposing questions. The questions consisted of a positive and a negative question described in section 4.1.3. The responses was categorised into various topics e.g. hoovering, cleaning, dish washing etc. was categorized as "cleaning". In addition other practical task like e.g. cooking food, walking the dog, trivial stuff, teaching, homework aid etc. was categorised as "other practical stuff". Using this method of categorising various topics into more covering categories, yielded six different categories, derived from 54 mentally fit participants divided between the different test conditions visible in the positive table 8.1 and the negative 8.2, and positiv/negativ responeses from the 12 mentally fit seniors that participated in experiment 2 visiblein table 8.3.

Positive Category	Senior PP & VP	Young PP & VP
Cleaning	30	31
Hygiene	13	0
Other Practical Stuff	21	32
Social Interaction	1	3
Unspecified Tasks	0	6
Robot Personality Traits	0	11

Table 8.1: Table showing the robot positive function categories derived from the after each test in experiments 1 young and senior participants. Physical Presence & Virtual Presence (is labelled as PP & VP) . The table also shows the number of times a categories was mentioned by the participants.

Negative Category	Senior PP & VP	Young PP & VP
Cleaning	1	0
Hygiene	8	0
Other Pratical Stuff	4	17
Social Interaction	5	6
Unspecified Tasks	0	0
Robot Personality Traits	0	30

TABLE 8.2: Table showing the robot negative function categories derived from the after each test in experiments 1 young and senior participants. Physical Presence & Virtual Presence (is labelled as PP & VP) . The table also shows the number of times a categories was mentioned by the participants.

Category	Positive Response	Negative Response
Cleaning	10	1
Hygiene	5	4
Other Pratical Stuff	23	6
Social Interaction	0	2
Unspecified Tasks	0	0
Robot Personality Traits	8	8

TABLE 8.3: Table showing the robot positive and negative function categories derived from the after each test in experiments 2 senior participants. The table also shows the number of times a categories was mentioned by the participants.

8.1 Analysis of Experiment 1 Interviews

The interviews provided various topics from the two age groups. Topics like cleaning, social interaction and other practical stuff was present in all experiment levels. The senior participant in both conditions listed "personal hygiene" as a category that robots should or should not be involved with. The young participants never mentioned hygiene as a topic that robots should be involved with. However the young participants formulated more personality traits, in both the virtual and physical presence condition to robots in generally compared to the seniors.

8.2 Analysis of Experiment 2 Interviews

The responses change with the robot in a context and now the focus of the senior participants is not cleaning or hygiene, but personal traits of the robot. The focus of these traits was not as such on that the robot should behave good or obey, like the young respondents wanted, but that it should act normal or social acceptable. Many of the senior participants also stated that they wanted the robot to be realistic in its movements and, that they would like to either look like a pet or a human.

8.3 Discussion of Experiment Interviews

The interviews yielded various different topics between the two age groups. Cleaning, social interaction and other practical stuff was present in all experiment levels. Both senior participants conditions listed "personal hygiene" as a category robots should/should not be involved with, and stated that would either in power the user to be more independent from others or to clinical and without the human touch. The young participants did not mentioned hygiene as a topic. However, young participants generally (both PP and VP) listed more personality traits to the robot than the seniors did, except in experiment 2 the senior started requesting personal traits of the robot. During the interview session the young participants seemed to be more uncomfortable with robots in generally then the senior participants, however they were still willing to let the robot teach, take care of children, and have a robot in their home and so on. All participants wanted the robot to do other practical work, which they thought could be nice to have a robot to perform for them. The interview at the end of the test sessions yielded a lot of information about what the different participants thought a robot should handle at home or in society, how it should behave and act towards people and what thoughts they had about robots. We are aware that the participants are biased by interacting with the robot, as the interview was conducted at the end of the test.

Discussion of Responses

Cleaning was common in both age groups and among the two levels of measure, and many of the participants stated that cleaning, was one thing they could do without and that it was a trivial task. The argument formed by the senior participants was that maybe health care staff would have time to be sociable, instead of cleaning the senior participants home. The young participants stated many times that they had other things to do then cleaning, so a robot would give them more free time. Hygiene was only mentioned a couple of times in the young participants groups and seemed not to be something they would need in their near future. However compared to a the senior participants, hygiene seemed to have a high priority. It could be that the senior participants are starting to feel the health effect of their age, and that they can see that they might need personal care in the near future. Many of them also mentioned that that a robot could in theory help them remain independent, and that they would not have to wait for health care staff to help them. The young participants groups weighted personal traits to the robot, as a high priority category. It was very important for the young participants that the robot obeyed and only did what it was told to do. Moreover these participants emphasised that the robot should not have the following traits or abilities; being evil, could run amok, be rebellious, able to take over the world, able to lurk around, be a smart mouth, commanding over the human user, calculated, have dreams, brush teeth on humans, teach, take care of children and so on. Even though the young participants had these traits and abilities of the robot as a major concern to them, they were still willing to have the robot in their home or let the robot take care of children in some cases. Some of the young participants told that they felt lonely when they got home, this could be why they were willing to let in a robot that they mistrusted into their home. Over all the young participants seemed to have been effected by movies like "I Robot" and "The Terminator" and often compared their own concern to movies like this. That the young participants did perceive anger in ABM1, could be effected by their suspiciousness of robots. Comparing these personality traits of the robot form the young participants. It is evident that senior participants responses change after interacting with the robot in experiment 2, and robot personal traits seems to be a subjects that senior participants thinks now, compared to hygiene which they now are not so focused on. The senior participants wanted that robot to look realistic and seemed like that they expected the robot to behave good towards humans, that it still should act in a socially acceptable manner. Overall it seems that, depending on what stage of life a person are in, the requirements and abilities of the robot varies. The elderly wanted hygiene and help to complete different task they thought was hard now, the young never thought of tasks like that, but were more concerned with the personality of the robot. Depending on what stage one is in their life and background one has, different tasks and abilities are thought of to the robot. This notion is strengthen by two participants whos age was between the two age groups, and were tradesmen. These middle age participants only wanted robots to lift heavy loads, do repetitive work or do dangerous work, as this was the normal kind of work the used to perform.

8.4 Robot Game Interview

Robots Gaming Skill

At the end of the interview the participants were asked if the robot was good at playing the game and, if they thought that the robot was god at guessing the outcome of the dice. One participant was undecided if the robot was good at playing the game or guessing the outcome of dice, therefore this participant's undecided reply was weighted as a no, hence 3 and not 2 replied no for both questions. Results for Robot Gaming Skills Eight answered that the robot was good at playing the game and three that it was not good at playing the game. Eight replied that the robot was good at guessing the outcome of the dice and three that it was not. The participants might have been biased in answering honestly, because it was a direct question in the interview and not a self-reported question on the questionnaire.

Chapter 9

Discussion

Results from experiment 1 suggest that semiors have reduced ability to recognize sadness of the intended movement expressed by a humanoid robot. Though differences could overall be grounded in that our experiment uses a different robot (Nao) compared to the robot (Robovie-X) Nomuras' experiment uses, which has different mass, dimensions and shape. However results presented in section 4.11 shows that seniors recognize happiness and anger, which is inconsistence with Nomura et al. [20] results for these expressed emotions. The difference could be explained by the different approaches to designing the robots affective body movement. Nomura's motions are based on Japanese puppet theater "Bunraku" whereas the ABM used on the robot in this project was a matched with full-bodied costumed actors. The ABM based on actors was verified by young viewers and so there is never created a baseline for when, and if, seniors actually do perceive anger through body movement.

Nomura compares which regions of the robot, the young and senior participants look at. Because Nomura uses self-reported measures to gather the data they could be considered to be more error prone as they relies on the subjectivity of the participant [28]. Therefore it is not possible to make a direct comparison with Nomuras results and the data gathers in this study, that is gathered in a objective way, as it relies on eye tracking equipment.

A statistical comparison between Nomuras and the results from experiment 1 of impression of motion speed and motion magnitude from senior participants, cannot be made without access to Nomuras raw data. However the reported findings suggest that

Japanese students and seniors have different impression of speed and magnitude. This is not the case for Danish participants who only differ in speed perception.

In our tests the variation in emotion identification can be explained with the multidimensional emotion representation. Certain emotions in the nominal representation of Ekman are closely related on the scales of arousal and affect. Using the multidimensional representation may have yielded more accurate results but the nominal method is simpler and more intuitive for participants.

9.1 Future Game Iterations

Several shortcomings were known before testing or became apparent during the final test which should be solved in a later version of the game. After the first guess in a round, the robot does not exclaim its guess when a face is detected (it waits for time-out). This increases the response time between responses but by solving the bug response time is decreased. The volume and quality of vocalizations should be increased as most participants had difficulties recognizing them. The dice recognition rate should be improved by adding several checks and comparisons between techniques to increase accuracy. The flexibility of dice detection should be improved by adding scan of the table surface, in addition with motion planning to ensure that the head camera is moved over the dice. In the current iteration of the test a non-dynamic animation placed the camera in the same position every time. A continous scan of the table should be implemented so the robot automatically reacts to every dice roll rather than relying on the user to trigger hand sensors before the robot looks at the dice.

Chapter 10

Conclusion

In section 3.2 a survey is conducted to ensure that ABM is successfully translated from actor to robot. The survey verifies that the movement transition from actor to robot is successful and also shows that Danish students perceive emotions through whole body movement of an actor and equally of a robot.

In experiment 1, section 4, Danish students and seniors are shows affective body movement of a robot and it shows that seniors are less capable at identifying sadness. Second iteration of experiment 1 changes the presence of the robot from co-located to virtual. This does not significantly influence ABM perception between students and seniors.

In experiment 2, section 6, contextualization of affective body movement did not improve emotion recognition in seniors. The results of experiment 2, section 6.2 shows that recognition of anger, happiness and sadness were not improved by showing emotions in the context of a game.

Cluster analysis in section 4.3 The interviews in chapter 8 yielded that young and old people want robots to conduct cleaning tasks for them. This was followed up by other practical stuff that the robot could do for the participants, however the young participants had some reservations on what practical stuff the robot could preformed for them. Seniors also had some reservations on the category hygiene, but the majority preferred robots aiding in hygiene tasks like e.g. toilet visits, and that a robot could make them independent of help form others. The young participants concerned about the personal traits and abilities of robots, many of the Young participants was very influenced by robot movies. Furthermore the young participants had very specific requirement for, what a robot could on its own, and that the robots abilities should not conflict with human existence. The senior participants did not have any requirements for the robots personally traits in experiment 1. In experiment 2 the senior participants had both positive and negative responses, however all the responses cantered on that the robot how the robot should act and look like.

The knowledge generated in this project is of significance to further interaction research and development of social robots.

Appendix A

Questionnaire

ABM emotion and magnitude/speed perception questionnaire used in experiment 1 and 2 seen on the next page.

(Der må gerne sættes flere krydser)
Robotten så ud til at vise:

Vrede	☐ Jeg er helt sikker på	☐ Det tror jeg	☐ Det tror jeg ikke
Had	☐ Jeg er helt sikker på	☐ Det tror jeg	☐ Det tror jeg ikke
Glæde	☐ Jeg er helt sikker på	☐ Det tror jeg	☐ Det tror jeg ikke
Overraskelse	☐ Jeg er helt sikker på	☐ Det tror jeg	☐ Det tror jeg ikke
Sorg	☐ Jeg er helt sikker på	☐ Det tror jeg	☐ Det tror jeg ikke
Frygt	☐ Jeg er helt sikker på	☐ Det tror jeg	☐ Det tror jeg ikke
Andet	☐ Jeg er helt sikker på	☐ Det tror jeg	☐ Det tror jeg ikke

Spørgsmål 2

Jeg synes hastigheden af robottens bevægelse var:

(Sæt ét kryds)

Meget langsom	○	○	○	○	○	Meget hurtig
	1	2	3	4	5	

Jeg synes størrelsen af robottens bevægelse var:

(Sæt ét kryds)

Meget lille	○	○	○	○	○	Meget stor
	1	2	3	4	5	

Appendix B

Questionnaire

Fixation questionnaire as as alternative if eye-tracker hardware should malfunction during testing, seen on next page.

(Der må gerne sættes flere krydser)
Robotten så ud til at vise:

Hoved. ☐ ☐ ☐

Jeg var meget opmærksom | Jeg var mindre opmærksom. | Jeg var ikke opmærksom

Arm. ☐ ☐ ☐

Jeg var meget opmærksom | Jeg var mindre opmærksom. | Jeg var ikke opmærksom

Hænder. ☐ ☐ ☐

Jeg var meget opmærksom | Jeg var mindre opmærksom. | Jeg var ikke opmærksom

Overkrop. ☐ ☐ ☐

Jeg var meget opmærksom | Jeg var mindre opmærksom. | Jeg var ikke opmærksom

Ben. ☐ ☐ ☐

Jeg var meget opmærksom | Jeg var mindre opmærksom. | Jeg var ikke opmærksom

Fødder. ☐ ☐ ☐

Jeg var meget opmærksom | Jeg var mindre opmærksom. | Jeg var ikke opmærksom

Andet ☐ ☐ ☐

Jeg var meget opmærksom | Jeg var mindre opmærksom. | Jeg var ikke opmærksom

Bibliography

[1] The role of physical embodiment of humanoid robot interaction: Focusing on backchannel head nods in danish first meeting encounters. 2.

[2] Hamid Aghajan, Juan Carlos Augusto, and Ramon Lopez-Cozar Delgado. Human-centric interfaces for ambient intelligence. page 544.

[3] Cynthia Breazeal. Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1-2):119–155, July 2003. ISSN 10715819. doi: 10.1016/S1071-5819(03)00018-1. URL http://linkinghub.elsevier.com/retrieve/pii/S1071581903000181.

[4] Peter Bruhm. *Forst Demens*. ISBN 87-614-0403-9.

[5] Mark Coulson. Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of nonverbal behavior*, 28(2):117139, 2004. URL http://link.springer.com/article/10.1023/B:JONB.0000023655.25550.be.

[6] Marco De Meijer. The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal behavior*, 13(4):247268, 1989. URL http://link.springer.com/article/10.1007/BF00990296.

[7] Kurt Demaagd. *Practical computer vision with SimpleCV*. O'Reilly Media, electronic edition, 2012. ISBN 9781449320362.

[8] Kirsten Sejerøe Szatkowski. *Demens - kommunikation og samarbejde*. Klim, 2002. ISBN 87-7955-145-9.

[9] Paul Ekman. *Emotions Revealed*. Times Book, Henry Holt and Company.

[10] Michael W Eysenck and Mark T Keane. *Cognitive psychology: a student's handbook.* Psychology Press. ISBN 9781841695402 1841695408.

[11] Jennifer Goetz, Sara Kiesler, and Aaron Powers. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003. The 12th IEEE International Workshop on*, page 5560, 2003. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1251796.

[12] M. Melissa Gross, Elizabeth A. Crane, and Barbara L. Fredrickson. Methodology for assessing bodily expression of emotion. *Journal of Nonverbal Behavior*, 34(4): 223–248, July 2010. ISSN 0191-5886, 1573-3653. doi: 10.1007/s10919-010-0094-x. URL http://link.springer.com/10.1007/s10919-010-0094-x.

[13] Theodore D Kemper. *A social interactional theory of emotions.* Wiley New York, 1978.

[14] Andrea Kleinsmith, P. Ravindra De Silva, and Nadia Bianchi-Berthouze. Recognizing emotion from postures: Cross-cultural differences in user modeling. In *User Modeling 2005*, page 5059. Springer, 2005. URL http://link.springer.com/chapter/10.1007/11527886_8.

[15] Mark L Knapp and Gerald R Miller. *Handbook of interpersonal communication.* Sage, Thousand Oaks, Calif., 1994. ISBN 0803948069 9780803948068.

[16] Rudolf Laban. *The Mastery of Movement.* March 2011. ISBN 978-1852731458.

[17] David McNeill. *Hand and Mind: What Gestures Reveal about Thought.* University of Chicago Press, August 1992. ISBN 0226561348.

[18] Melvin J. Millon, Theodore. Lerner. *Handbook of Psychology.* John Wiley & Sons, Inc.m Hoboken, New Jersey, 2003.

[19] Toru Nakata, Tomomasa Sato, Taketoshi Mori, and Hiroshi Mizoguchi. Expression of emotion and intention by robot body movement. In *Proceedings of the 5th International Conference on Autonomous Systems*, 1998. URL http://staff.aist.go.jp/toru-nakata/IAS.pdf.

[20] Tatsuya Nomura and Akira Nakao. Comparison on identification of affective body motions by robots between elder people and university students: A case study in

japan. *International Journal of Social Robotics*, 2(2):147–157, 2010. ISSN 1875-4791, 1875-4805. doi: 10.1007/s12369-010-0050-2. URL http://link.springer.com/10.1007/s12369-010-0050-2.

[21] Robert Plutchik. *Handbook of cognition and emotion.*

[22] Byron Reeves. *The media equation : how people treat computers, televisions, and new media as real people and places / Byron Reeves and Clifford Nass.* Cambridge University Press, Cambridge, 1996. ISBN 157586052X.

[23] Culture-Aware Robotics CARsMatthias Rehm. Experimental designs for cross-cultural interactions: A case study on affective body movements for HRI. 2013. URL http://www.create.aau.dk/irc/culture-aware-robotics-cars-matthias-rehm/.

[24] Matthias Rehm. From multicultural agents to culture-aware robots. In Masaaki Kurosu, editor, *Human-Computer Interaction. Human-Centred Design Approaches, Methods, Tools, and Environments*, number 8004 in Lecture Notes in Computer Science. Springer Berlin Heidelberg, jan 2013. ISBN 978-3-642-39231-3, 978-3-642-39232-0.

[25] Alvin C. Rencher. *Methods of Multivariated Analysis.* 2002.

[26] Klaus R. Scherer and Paul Ekman. *Approaches to Emotion.* L. Erlbaum Associates. ISBN 0898593506.

[27] Inc Sight Machine. Simplecv @ONLINE, May 2014. URL http://simplecv.org/docs.

[28] Tom Snowden, Robert. Troscianko and Peter Thompson. *Basic Vision.* Oxford University Press UK, 2006.

[29] Kazuhiko Takahashi, Mana Hosokawa, and Masafumi Hashimoto. Remarks on designing of emotional movement for simple communication robot. In *Industrial Technology (ICIT), 2010 IEEE International Conference on*, page 585590, 2010. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5472735.

[30] Joshua Wainer, David J. Feil-Seifer, Dylan A. Shell, and Maja J. Mataric. The role of physical embodiment in human-robot interaction. In *Robot and Human*

Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on, page 117122, 2006. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4107795.

[31] Harald G. Wallbott. Bodily expression of emotion. *European journal of social psychology*, 28(6):879896, 1998. URL ftp://ftp.infomus.org/pub/PPM/Publications/Wallbott,%201998,%20bodiliy%20espression%20of%20emotion.pdf.

[32] Bonnie Wong, Alice Cronin-Golomb, and Sandy Neargarder. Patterns of visual scanning as predictors of emotion identification in normal aging. *Neuropsychology*, 19(6):739–749, 2005. ISSN 0894-4105. doi: 10.1037/0894-4105.19.6.739. URL http://psycnet.apa.org/journals/neu/19/6/739.html.