Institute for Visualization and Interactive Systems

University of Stuttgart
Universitätsstraße 38
D–70569 Stuttgart

Bachelor’s Thesis Nr. 272

Observation of Human Behaviour in Reference to Human-like Motions of Animated Avatars

Sascha Sprott

Course of Study:	Informatik
Examiner:	Jun.-Prof. Niels Henze
Supervisor:	Dipl.-Ing. (FH) Valentin Schwind
Commenced:	November 11, 2015
Completed:	May 10, 2016
CR-Classification:	H.1.2

Kurzfassung

Über Moris Theorie des “Uncanny Valleys” - der Effekt, wenn Menschenähnlichkeit und Unheimlichkeit für das Erscheinen von Robotern und virtuellen Avataren, in einem nicht-linearen Zusammenhang stehen - existieren viele unterschiedliche Meinungen und unterschiedlichen Untersuchungen. Einige haben die Beobachtung gemacht, bewegte Avatare sind nicht mehr oder weniger sympathisch als unbewegte. Allerdings wird oft impliziert, dass ebenso Interaktionen keinen Effekt zeigen. In dieser Thesis wird untersucht, inwiefern sich das menschliche Verhalten gegenüber computergenerierten Avataren verändern, wenn eine Interaktion stattfindet. Dafür wurden echte menschliche Bewegungszüge aufgenommen und auf virtuelle Gesichter übertragen. Dann wurde eine Studie mit 18 Teilnehmern durchgeführt, in der jeder vier verschiedene computergenerierte Avatare zu sehen bekommen hat. Dabei haben die Teilnehmer in zufälliger Reihenfolge erst eines der Bild einer Figur gesehen und danach mit dieser in animierter Form interagiert. Die Interaktion wurde durch ein Spiel “WER BIN ICH?” realisiert. Dabei wurden die Probanden in einen bekannten Charakter versetzt und mussten diesen durch “Ja/Nein” Fragen erraten. Nach jedem Bild und jeder Interaktion haben die Probanden einen Fragebogen zur Bewertung von Menschlichkeit, Unheimlichkeit und Attraktivität von virtuellen Figuren und Robotern ausgefüllt. Daraus haben sich Ergebnisse evaluieren lassen, welche zeigen, dass Interaktion zumindest die Menschenähnlichkeit und die Attraktivität von virtuellen Figuren steigert.

Abstract

There exist many different opinions and miscellaneous researches regarding to Moris theory about the “Uncanny Valley” - meaning the effect, when human likeness and eeriness stand in a non-linear relationship for the appearance of robots and virtual faces. Some of them depict that motion do not affect the valley in any way, but they implied it wrongly for interactions. In this thesis we wanted to investigate in the change of behaviour, when humans interact with virtual 3D avatars. Therefore we recorded various animation of a real human and mapped them onto virtual faces. Then we designed an experiment with 18 participants, who were shown four different inanimated computer generated faces and afterwards the same faces animated for an interaction, in a random order. The interaction was made with a game called “WHO AM I?”, where they slip into the role of a well known character and than had to figure out through “yes” or “no” questions who they are. After each inanimated face as well as after each interaction they had to fulfil a questionnaire, to rate human likeness, eeriness and attractiveness of the virtual avatars. We evaluated significant positive effects of human likeness and attractiveness due to interactions with computer generated avatars.

Contents

List of Figures

List of Tables

Acknowledgement

The Author of this thesis would like to thank the actress Deborah Freudigmann, without whom the experiment would not have been possible and who agreed for participation in the short term. Further thanks goes to the participants of the user study and substitutes who found, also in the short-term, time for participating, and to the anonymous reviewers. Special thanks go to the examiner Jun.-Prof. Henze, who awarded this topic to me, and my supervisor Valentin Schwind for the cooperation and help he provided to complete this thesis.

1 Introduction

1.1 The Uncanny Valley

In 1970 Masahiro Mori, a Japanese robotics professor, hypothesized an odd effect in human-robot-interaction. This effect describes the phenomenon, when increasing robots human-like visual appearance, their familiarity grows too, but when reaching a particular level of similarity, the familiarity turns into eeriness. After that valley, when the non-human entity turns into a real human-like, the familiarity reaches its maximum. Mori called this anomaly “The Uncanny Valley” (UV)[1]. For the following years, this discovery did not receive much attention. Later, a few years ago, when robots and also computer animated avatars became more and more present in everyday life, scientists revived this hypothesis. Research reached from biological investigation in the origin of the UV, over more human-robot-interaction, to investigation in computer generated virtual figures.

For Mori, the purpose of robots have a high impact on their familiarity. Industrial robots are made to replace humans, but do not need to copy visual human likeness. In contrast, the visual design of toy robots is more conscious. Since they might have human-like extremities, eyes and a mouth, babies feel attracted to these kind of toys. When continuing to increase human likeness, it comes to a point where robots look human. For partial instance, prosthesis are developed to replace real human body parts. Therefore they do arouse affinity in us. Nowadays these prosthesis are barely distinguishable from real body parts, whereby some years ago they would look *creepy* and *dead*. Even if not, when we know about its artificiality, we experience uneasiness.

Because movement is a part of living creatures, Mori says it will “amplify the peaks and valleys” from the UV Graph (Fig. 1.1). For illustration, switched off industrial robots are just machines, but when switched on and when imitate human activities, we feel some empathy.

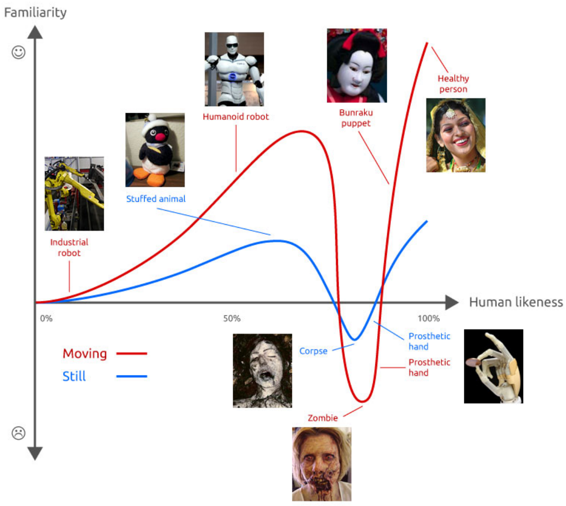


Figure 1.1: The uncanny valley effect, discovered by Mori[1]

1.1.1 Explain the Uncanny Valley

Threat avoidance

When it comes to reconnaissance of the UV, several theories exist to explain its origin. Mori assumed some kind of instinct for self-preservation[1]. As some scientist explored in studies, "the uncanny valley is the result of an evolved mechanism for pathogen avoidance"[2]. An organism or figure with human-like features and looking, may appear diseased. Hence human beings perceive the feeling, the opposite organism could carry diseases. To protect themselves, human than respond with disgust and refrain from the organism analogical figures. Further expositions associated with self-preservation could be a "fight-or-flight"[2] reaction. Therewith, fear would be another possible explanation as originator for the uncanny valley.

Similar brain-circuits activities

Another approach is made over shared circuits in our brain. It is about the presumption, that perceptual, cognitive and affective processing work in concert[2] when human perceive eerie gestalts. The point is that, human-like beings such as actual humans, robots or animated avatars, activate these circuits when human perform intentional actions and when they watch another entity performing similar motions. For example Jabbi, Bastaansen and Keysers[3] found out, that the experience of disgust, and the bare sight of someone else experiencing it, are closely linked in our brain. But robots do not stimulate these same regions and therefore do not trigger the same reactions.

Predictive coding

Is about that brain activity will increase for stimuli that are “not well-predicted or explained by a generative neural model of the external cause for sensory states”[4]. Human beings evolve with their experience and currently human associate human appearance with biological motion and robots with mechanical motion. Therefore humans predict this exact thing to happen. If not, a prediction error occurs[4]. This processing conflict need to be resolved by the neural network.

Categorical perception

Means the phenomenon where the perception of someone is altered through the categories he is used to know[5]. In particular, someone is more able to make perceptual differentiations about things, when those things belong to adjacent stimuli categories and a boundary is drawn between these. For equal pairs of stimuli this effect is neutralized.

Evolutionary aesthetics

Attractiveness differs from one culture to another. There are high agreements on what is attractive, but each culture favor unique features of their own society. “The perception of attraction has a biological basis” [2] and hence is unconscious automatism. Humans tend to partners with “fertility, healthy and social desirability”[2].

Terror Management

Commonly used as explanation, this declaration is about subliminal reminders that lead to a “shift in our attitudes and preferences”[2]. It is about our preference of well known things. We accept and believe in the cultural values we know. If someone or something is questioning these qualities, we are not prone to change our worldview[2].

1.2 Animated Avatars and their Significance in our life

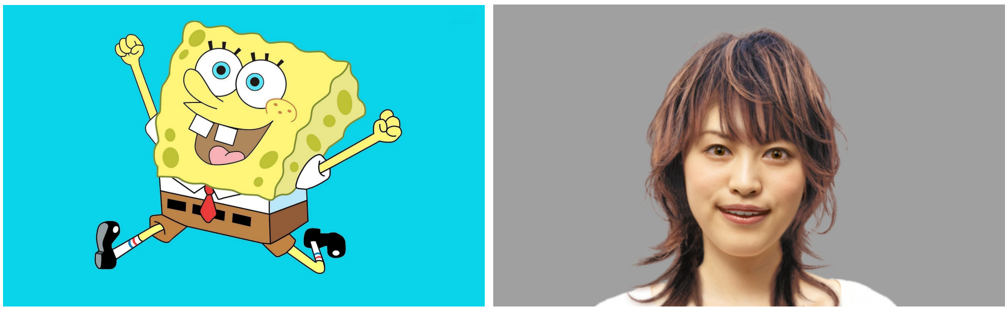
These days robots and especially animated avatars are found in many sections of our lives. Everyone in our society is watching movies, playing computer games or mobile games on Smart-phones. Even if we do not want to use computers, the commercials on streets or TV’s contain computer generated (CG) avatars. There are not many more opportunities left to avoid computer animated figures. With the advancing technologies CG avatars are seen more frequent. Therefore it has a high importance, that these avatars do not trigger uncomfortable feelings in us if we do not intended so. They should look neutral or at least make us feel good or attracting to them so that we are not negative affected. For example figures like *Davy Jones* from *Pirates of the Caribbean* or *Gollum* from *Lord of the Rings* as shown in Fig. 1.2a and Fig. 1.2b, are intended to look creepy. Hence we feel disgusted and reluctant. On the other side, figures like the conductor from *The Polar Express* seen in Fig. 1.2c are ment to look human and likeable but make us feel weird when we look at them. They show us human-like motions and a visual human analogy, but not completely human-like. A similar example is the movie *Mars Needs Moms*, where human-like characters like Milo (Fig. 1.2d) are shown, who are intended to look human. These figures trigger discrepancies to our awareness of humans.

At first sight this indisposition, which arises when seeing or interacting with figures like *Milo*, is the only consequence that occur. When we look at the achievements of films with CG characters, we see significant differences in profit. The movie *Mars Needs Mums* called for a budget of $150 million but had a total gross about $21 million[10]. Unlike the movie *The Incredibles*, a movie with human-like figures too, but such with no intend to look perfectly human. It had a budget of $92 million but a total gross

(a) Davy Jones[6] **(b)** Gollum[7]

(c) Conductor[8] **(d)** Milo[9]

Figure 1.2: Intended uncanny designed avatars



(a) Spongebob SquarePants[11] **(b)** Ucanny Girl[12]

Figure 1.3: Unreal character versus uncanny character. Examples of CG characters who do or do not fall into the uncanny valley.

about $261 million[10]. This extreme difference may come from a better respectively worse story-line, but also from the fact, that the characters in *Mars Needs Mums* fall into the UV. According to the fact, the UV has a serious effect on our economy, the further investigation in this topic is of high importance.

1.3 Research Question

So far we know, the appearance of CG avatars and robots has an impact on how we perceive these. But what about interactions. Maybe an interaction change the way we perceive living or feign living entities. Therefore the interaction between human and 3D animated avatars is crucial and research is a must-do. Thus, we strive investigation with avatars which originate from different levels on the uncanny valley curve. By this means we can measure differences in human behaviour between the different interactions with CG faces. Heeding Mori's theory, when interacting with obvious unreal looking avatars, like Spongebob Squarepants (Fig. 1.3a) we should not feel uncomfortable, but instead exhilarated. On the other hand, interactions with the figure shown in Fig. 1.3b should be pretty strange (An interaction is possible at `http://www.cubo.cc/creepygirl/`) because this figure falls into the uncanny valley. Thus there can be evaluated a differences in human responsive behaviour for diverse figures from various levels out of the uncanny valley graph.

We need a more conscious approach in the development of CG avatars and design of human-like robots. Therefore the investigation on different artworks of uncanny or not uncanny avatars is of great importance. The current research in this thematic area exist, but is lucid and inconsistent. Most of it is done in human-robot-interaction, where robots with different appearances - taken from varying levels of the uncanny valley - interacted with human. Humans than made statements about their feelings and experience with these robots. There exist research in human-avatar-interaction with computer generated characters. Unfortunately the most interactions were not real, but only videos which showed CG avatars. Mori write primary about static robots, but pointed out the amplification of the uncanny valley effect through motion, because movement imply vividness. We do not know neither, if and how the UV influences interactions. Hence, investigation for human-avatar-interaction with 3D animated faces is a gap in science.

2 Related Work

2.1 About Computer Generated Characters

For the investigation in this thesis, we worked with mapping the motions of a human out of real space into virtual space. This mapping should for sure be as veridical as possible, in order to guarantee an accurate interaction. Francesco Cafaro investigated in this topic[13], with a study where users had to rate their body movements (a jump). The motion was captured through a system called *CoCensus*. This system visualizes data in real-time. The users had to jump 12 times, each time the jump was replicated by the system, but with different conditions. The system had to manipulate three factors of the motion: The sensitivity to the nuance, the synchronism and the physical realism. The outcome of this survey depict, that if one of the mentioned factors for veridicality was manipulated users perception indicated a valley, like the UV. Thus, when designing CG avatars and a limitation of one dimension is not suppressible, a mixture of limitations of all dimensions is recommend.

2.2 Virtual Figures and the Uncanny Valley

The examination on the UV has led to many different results. Tyler Burleigh[14] investigated in 2009, for example, the existence of the UV through an empirical study. Therefore 164 participants where led to judge the appearance of CG faces. The measurement parameters were *human likeness, eeriness, scariness, disgust* and *attractiveness*. The evaluation of the data showed “linear relationships...between human likeness and all emotional responses”[14]. Controversial to the UV, where a non-linear relationship should be at least between human likeness and eeriness, he observed a linear relation between human likeness and the other four different feelings. As expected, eeriness was “found to be most strongly related to fear followed by disgust”[14]. Burleigh concludes that the appearance of eeriness is a consequence of “threat avoidance or terror management process”[14], which might be a correct resolution.

Another study from Dr. Edward Schneider, Yifan Wang and Shanshan Yang [15], about the occurrence of the UV with video game characters validated Mori's theory about robots for CG avatars too. Their study worked on the relationship between human likeness and attraction from CG avatars. It was executed with 60 participants, which were asked about their opinions of 75 CG avatars from inside and outside video games. The evaluation replicated Mori's statement and prove the existence of the UC effect for virtual characters[15]. With that knowledge further investigation in human-avatar-interaction is feasible.

While the emergence of the UV in CG avatars is elaborated, Laura M. Flach and Vanderson Dill[16] took it to the next step and conducted a survey with the main question: "How do people perceive Computer Graphics made characters?"[16]. The study was conducted online and reached a total of 210 responses. The participants where asked to answer questions about 17 different CG avatars, taken from different levels of the UV. In the first part the participant saw just images of virtual figures, in the second part each figure was animated and in motion. As expected, people responded at least partly with Moris theory and assumption about motion. As the evaluated data shows in Fig. 2.1, motion deepened the Valley but no noticeable differences on other locations of the Graph were observed. Unfortunately there were no real interactions between subjects and figures, because they only were asked about their opinions and could only see but not interact with the figures.

The closest conjunction to what this thesis is about is made by Rachel McDonnell and Martin Breidt. In their survey "Face Reality: Investigating the Uncanny Valley for virtual faces"[17], 24 subjects viewed 108 movies in which a CG avatar made statements. Afterwards the participants were asked to tell if the model just told the truth or lied. The model was generated out of a real woman. The series of statements was recorded on the real model and with some further photos taken, a virtual model was created, matching the original. To generate real gestures, the model was asked questions from an interviewer and was told whether to lie or tell the truth. The voice of the actor was recorded in original quality. Furthermore the model was rendered in three different modes: *High Quality, Game Quality* and *Non Photorealistic (NPR)*. The results showed, that statements from all different rendered models were graduated as equivalent. There were only a small deviation, that the statements of the HQ face where rated more often as lies. But for the Game face or the NPR no such differences were found. According to those findings, CG avatars rendered in HQ and thus more human likeness, let us prone to mistrust them. Hence we have a shift in behaviour when interacting with CG characters. But the limitation of this survey is, only the quality was changed, not exactly the human likeness.

As the revisited elaborations prove the existence of the UV, there also exist evidence from measured electrical brain activity. Burcu A. Urgen investigated neural systems in the

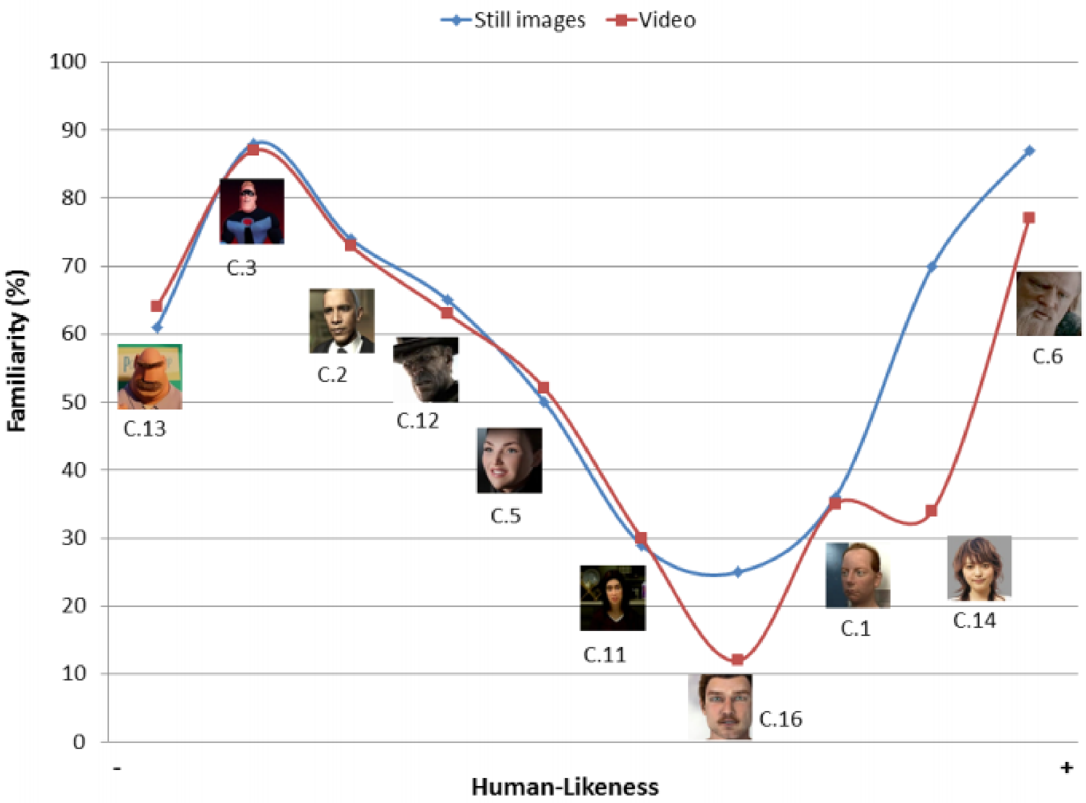


Figure 2.1: Results of Laura M. Flach and Vanderson Dills survey[16]

brain based on the explanation *Predictive coding theory*[18]. As the brain needs to resolve conflicts its activity rises and can be measured. He conducted a cognitive neuroscience study, where the brain activity was measured using electroencephalography. Participants of this survey were requested to view three different avatars. First a human, second a robot with similarity to the human and third and last a full obvious robot. The human had human motions and the robot mechanical motion. The human like robot however, had a mismatch in appearance (human) and movement (mechanical). As expected, the human like robot violated “at least partially the brain’s internal predictions”[18].

So far we know the existence of the UV is real, but the influence of motion is still unclear, as we could only determine an amplification in eeriness. However there are studies who did not discover such an amplification as Mori predicted in his theory. Lukas Piwek, Lawrie S. McKay and Frank E. Pollick evaluated in a study [19] no amplification for the UV in motion. The designed full-body CG avatars and applied an adjustable motion set to these avatars. Thereby the motion of the avatar could be changed respectively improved. Their discovery is in conflict of what Mori said. They did not just discover no amplification, but an increase of familiarity in the UV due to improvement of motion.

Furthermore another study investigated the same topic and hit upon similar results. James Thompson designed a study with colleagues where two full-body CG avatars were used for investigation[20], a mannequin and a human-like figure. The movements were recorded from a human actor. To scale the motions different values were manipulated, including the joint articulation of arms and legs, the phase synchronism of the arms respectively legs and random generated non-human jerks in motion. For both figures and through all of the three different manipulations, the familiarity increased linearly with increasing human-likeness. Therewith we have a high and crucial inconsistency with Mori's theory about the effect of motion. More likely these findings represent the theory of motor resonance, which states that with increasing robots human like movement, they begin to match with human motions and hence become familiar to humans.

2.3 Human-Robot Interaction

The previous research in human-avatar-interaction is quiet rare. Thus we contemplated work about human-robot-interaction, so we can learn and maybe transmit some of the findings.

In 2009 Christian Becker-Asano conducted an experiment on the ARS Electronica festival[21]. The survey included the analysis of the interactions between 24 visitors and the robot Geminoid HI-1 (Fig. 2.2a right). The participants then had to held a conversation with Geminoid HI-1. Subsequent they were asked about their feelings during the interaction. The evaluation led to a surprising result. Only 9 participants felt uncanny during the conversation, 7 felt odd and 5 changed their feelings during the interaction. The main factor for uncomfortable feelings were the unpredictable motions and expressions of Geminoid HI-1. The subjects did not feel dominant in the conversation due to the fact that the robot produced inappropriate social signals. This experiment shows us furthermore, that uncanny is not uncanny. Some people might have other perceptions of interactions with robots, than what we expect according to Mori.

Another study, where Geminoid HI-1 participated with another robot(Fig. 2.2b) too, was carried out by Jakub A. Złotowski[22]. Participants of this study had three interactions with each robot. The first had the purpose that the participant should become familiar with the robot. In the second and third interaction the subjects had to apply to a job and therefore convince the robot. As measurement the study used likeability and eeriness. The main findings that were worked out, indicates that embodiment as well as attitude affects likeability and eeriness through repeated interactions.

(a) Geminoid HI-1 with its creator Hiroshi Ishiguro [21]

(b) Robovie R2 [22]

Figure 2.2: Human-like robot vs on-human-like robot

Another aspect of this empirical study, that is in our interest, is the structure and process. After each interaction the participants had to fill out a questionnaire. The participant was alone through the whole study and only had interactions with the robot. The robots were remote-controlled from a "Wizard-of-Oz room". During the completion of the questionnaire, participants had no visual contact to the robots. The experiment of this thesis will be oriented towards Jakubi's study.

Further from interest for us, is an evaluation from Chin-Chang Ho, Karl F. MacDorman and Z.A. Di Pramono [23]. In a study they investigated in ratings of 18 figures (Fig. 2.3) with 27 emotions. The figures' visual appearances differed in level of human likeness along the UV graph, starting with a 100% non-human-like robot as a baseline for familiarity. There were five main findings of the study. First, uncanny robots are described best through the emotional characterizations *eerie* and *creepy*. Second, the descriptions *eerie* and *creepy* are primary related to *fear, shocked, disgust* and *nervous*. Third, the description *strange* is little associated with emotion and therefore it may be more cognitive. Fourth, human features increase human-likeness and fifth, woman sensitivity towards eerie and creepy is higher than men and older men are prone to see human likeness in robots in spite of their eeriness.

2.4 Measure of human likeness

The presented related work had one common aspect, they all had to measure human likeness. In other words, the relation of CG virtual avatars and robots to human affinities.

Figure 2.3: 1 Human being and 17 Robots of different levles of the UV[23]

But only ratings of interactions are not enough. There is a need of standardization of measurements, such that results can be compared and brought into context. Therefore Christoph Bartneck brought together five key concepts in human-robot-interaction[24]. These concepts are: *Anthropomorphism, Animacy, Likeability, Perceived Intelligence* and *Perceived Safety*. For each concept he worked out a questionnaire, where ratings could be done through differential scales. This collection of questions is called "Godspeed"[24] and is commonly used as a rating system for human-like characters out of human-robot/avatar-interaction.

Chin-Chang Ho and Karl F. MacDorman assumed[25] that the index construction of the UV may not be correct, because of their strong correlation. This could lead to an inaccurate plot of CG/robot characters. Therefore they developed and validated an alternative to the Godspeed indices. To accomplish this, they first tested the original Godspeed with the close correlated indices and second developed new indices with less correlation. The new collection of questions simplify plotting relations for characters with varying human likeness.

2.5 Monkeys and the Uncanny Valley

Due to the fact, that computer animated human like avatars or robots fall into the UV effect, Shawn A. Steckenfinger and Asif A. Ghazanfar conducted a study, where monkeys

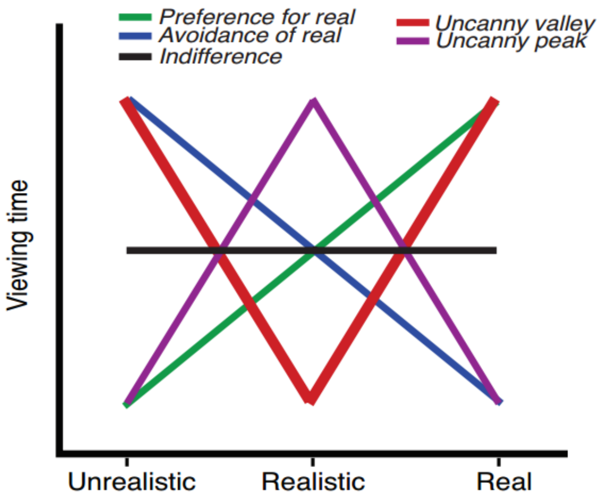


Figure 2.4: Possible monkey's behaviour[26]

on different computer generated faces were observed[26]. Monkeys are animals with the most human like behaviour. Their behaviour is traceable on primary instincts and not falsified. Furthermore from human point of view, the UV is not what we expect to happen. This could be because of our evolution and familiar life with robots and animated avatars. With monkeys and their primary instincts, the origin of the UV effect could be revealed. Therefore Shawn A. Steckenfinger and Asif A. Ghazanfar selected five macaque monkeys (*Macaca fasicularis*) as their subjects. Then they showed real monkey faces (real), synthetic agent faces (realistic) and unrealistic synthetic faces (unrealistic) to these monkeys. The faces had three different types of expressions: "coo", "scream" and a "neutral" face, presented in static and animated form. For each section one face was shown. Then the time was measured, how long the monkeys looked at one faces, but also the number of fixations to a face made by the monkey. This was done due to the assumption humans tend to look longer and more often at pleasant faces than uncanny faces.

The Expectation of this study was, if monkeys fall into the UV effect as humans do, then they should "prefer to look at unrealistic synthetic faces and real faces more than to the realistic synthetic agent faces" [26]. Five different outcomes where presented in the study, which could have arise (2.4). They could show similar behaviour to each face (black line), since they looked the same amount of times and same period at each face. Another outcome could be a longer fixation period and a higher amount of fixations for the unrealistic face and a shorter period and fewer amount of fixations for the real face (blue line) or the other way, preference of the real face and avoidance of the unrealistic face (green line). The last two outcomes could be an uncanny peak (purple line) or an UV (red line).

Eventually, even though there were five different possibilities, all monkeys revealed behaviour which falls into the UV effect. The difference between the amounts of fixations and the period of fixation from unrealistic and realistic faces was significant. The same holds for the difference between the realistic and real faces. No such difference was detected between unrealistic and real faces. Furthermore, analogical observations were made for animated faces.

Summarizing we can say, monkeys visual behaviour matches with humans' and falls into the UV effect. This means humans are not the only life form with such indisposition towards realistic looking but not real conspecifics.

Unfortunately there where some limitations in this study. In fact Shawn A. Steckenfinger and Asif A. Ghazanfar could not determine if the subjects find realistic faces more or less attractive than the unrealistic or real face. Neither could they make statements about the monkeys "experiencing disgust or fear"[26]. These restrictions are steered by the fact that the study operated with eyetracking only. But as monkeys are related to humans, Shawn A. Steckenfinger and Asif A. Ghazanfar expected them to show similar behaviour and therefore that they show some of the same emotions. As future work, Shawn A. Steckenfinger and Asif A. Ghazanfar mentioned pairing of synthetic faces with real voices, and work out if humans or monkeys provide the same results[26].

3 Approach

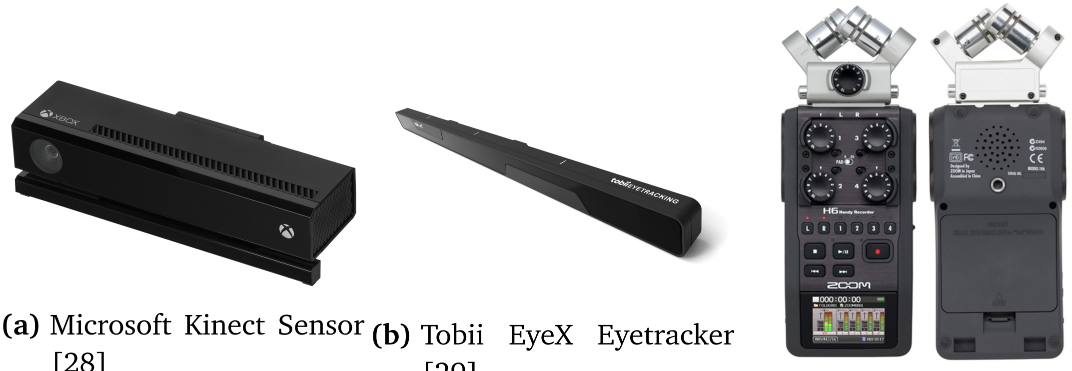
3.1 Method

3.1.1 Performance Capturing

The experiment related to this thesis is designed to measure change in human perceptual behaviour, during the interaction with CG avatars. These figures are in human-like shape and should therefore, if possible, match human in all features, also in motion. To achieve the optimal translation from motions out of real space into motions in virtual space, we used the technology of optical performance capturing. This technique records movement of people or objects and translate motions into a data stream. In the early days motion capturing was done by taking multiple images of objects and then calculated their position in space. There exist various methods to track motions in reality and map them on CG entities. To give a quick overview, a few different methodologies will be presented.

Starting with optical systems that uses markers as reference, there are again various different methods. Active markers are for instance LEDs which are illuminated and then tracked by the system. Passive markers on the other hand, are retroreflective markers, like those used for the OptiTrack System[27]. However, markerless systems do not force users to wear some kind of special clothes or special objects. They identify shapes and their track movements. There exist also non-optical systems too, but as we use an optical system in the experiment of this thesis, deepening in this thematic would go beyond the scope.

For our experiment we used an optical tracking system. In fact we used a markerless system, a *Microsoft Kinect Sensor*(Fig. 3.1a). The Kinect contains an RGB camera and a depth sensor. With these components a full-body tracking in 3D space is possible. The sensor provides further features, like face and voice recognition. With the infrared laser, which is integrated in the depth camera, video data can be captured in any light conditions. We used this type of motion capturing, because it can be handled easy and no special features need to be considered. The only limitation was, the user has to stay



(a) Microsoft Kinect Sensor [28]

(b) Tobii EyeX Eyetracker [29]

(c) Zoom H6 Audio-Recorder [30]

Figure 3.1: Capturing Apparatus Hardware

in a particular capturing-box. If the user is to close to the sensor, motions get blurred and the Kinect sensor loses track.

3.2 Apparatus

For this thesis, we had to develop two different programs. As we had to record all necessary human-motion data in the first step and in the second step replay and assign them in another application to virtual figures.

3.2.1 Capturing Apparatus

To process the data, captured by the Kinect, we took another application: *Brekel Pro Face*, a motion capturing Software. Pro Face make it possible to record and stream 3D face-tracking data for up to six faces[31]. For each face, Brekel is able to track 21 face morphings, also called face animation units, which are responsible for facial gestures and expressions: *Brow Up L, Brow Up R, Brow Down L, Brow Up R, Eye Closed L, Eye Closed R, Cheek Puffed L, Cheek Puffed R, Lips Pucker, Lip Stretch L, Lip Stretch R, Lip Lower Down L, Lip Lower Down R, Smile L, Smile R, Frown L, Frown R, Jaw L, Jaw R, Jaw Open* and *head* respectively *neck* position and rotation. Each of the morphings is transmitted as a value between 0 and 1, where 0 means neutral position and 1 maximum facial expression. Neck and head rotations are indicated in Euler-angles. The neutral position of a face can be set in Brekel, when a person is certain about not making any facial expressions and is looking completely neutral. The position can be stored and Brekel will compose the facial expression thenceforth. Brekel also was able to calibrate an user, by tracking the face when the user performed a sequence of extreme facial expressions. Through the calibration, Brekel figured the maximum and minimum morphing values for each animation shape.

Further we recorded Eye-tracking data, because Humans attention is noticeable through eyes and the virtual figure should not appear careless or stiffen due to static eyes. Neither with random generated eye movements, the recreation of the situation in virtual space, as it was in reality, is possible, since random gazes would not reflect human likeness. To track the eyes we used the *tobiiEyeX* Eye-tracker[32] (Fig. 3.1b). This tracker is capable of many different recording options. The eye-tracker uses two different coordinate systems, one for the positioning the eyes in space, in front of the sensor and another to map and track gaze-points on a desktop. Thereby the EyeX Engine API is possible to locate the position of the eyeballs in real space and convert it into virtual space position. Furthermore it can calculate the gaze-point for gaze-beams. That means the point on desktop screen where the user is directly looking at. While these gaze-points change very quickly and would result in a very rapid and volatile movements of the eyeballs, the API provides “fixations”. Fixations are sets of gaze-points over a specific area, collected over a brief amount of time. With the help of these features, the Tobii EyeX enables some interacting features: gaze-aware region, activatable region, pannable region and user presence[32]. We will not discuss these features here, because they were not a part of this thesis. But for the record, we stored all possible data of the Eye-tracker. The Eye-tracker API provided, just like the Brekel software did, a calibration tool, such that there are as little as possible deviations and incorrect measurements.

As last Part of the capturing apparatus, we recorded sound. Therefore we used a portable *Zoom H6 Microphone audio recorder* with an XYH-6 X/Y capsule (Fig. 3.1c). The mic provided a direct USB connection, such that we could start the record simultaneously with the face and eye data. The Microphone was accessible through a simple script in Unity, where connected Microphones could be called.

To store these data, we developed a program using the *Unity Game Engine*. Via a TCP Stream we sent the face data from Brekel to Unity. The face morphings, neck and head values then where assigned to a base face(3.2). The face then imitated the same motions as the face in the Brekel software and the user in front of the Kinect Sensor. The Eye-Tracking Data could be accessed via a script and was directly stored, without assignment earlier, since the base face had no eyes. We developed the program, such that the recording was asynchronous. That means we did not use the Unity *Update*-Method for storing the data stream, because the amount of calls were too little. The text file then had all together 48 entries for each line of Data recorded. The sound files had a bit rate of 705 kb per second and were recorded in stereo. Therefore, for each text file, there existed a sound file too.

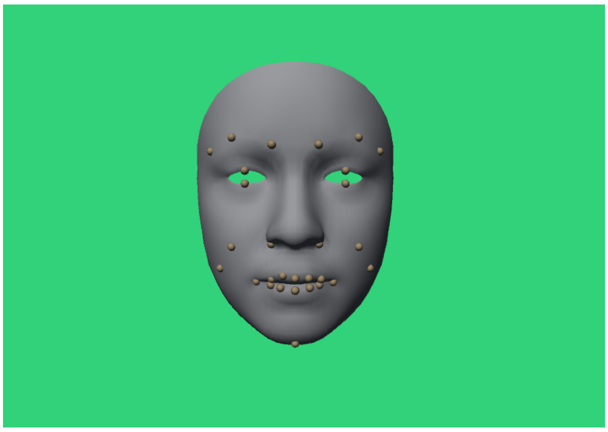


Figure 3.2: Base face, used in Unity to visualize morphings and neck/head rotation

In total we recorded 128 animations, from which we used only 61 for our experiment, because the rest were broken or could not be used due to various reasons.

3.2.2 Interacting Apparatus

For the experiment to be conducted we developed another program with the *Unity Game Engine*, to replay the recorded data of the sound and text files, record personal data of the probands and fulfil and safe results of the questionnaire. The application contained ten different scenes, one questionnaire-scene, four scenes for the animated avatars, four to show the inanimated images of the avatars and the last one as a "Goodbye"-scene. We constructed the program, such that after each avatar-scene, it switched back to the questionnaire-scene. Hence subjects could answer the questionnaire after each figure, inanimated or animated. The questionnaire, composed out of the *Godspeed indices*, had a scale for each question and was completely in English. This scale delivered a value between 0 and 1. At the end of each questionnaire all data were written into a result-file in CSV-format. Thus a file contained at the end the personal data of the subject and the values from the questionnaire for each avatar-scene, thereby eight outcome-lines.

For the animated avatar-scenes, we had to reload the stored audio and text files. The sound files were loaded with the *WWW Scripting-API*[33] which Unity provides actually to download content from websites or servers. We used it to load all used sound files at the startup of the program, to make them available the complete runtime without any more loading times. To play audio in Unity, an *AudioSource* hast to be created and assigned to a *GameObject*, where an *AudioClip* can be assigned and later be played. While the sound files were be loaded at start up, we just stored the paths for the text files then. The application consisted of two main functionalities, realized through Unity coroutines. These are available when declaring the Methods with an *IEnumerator* return type. A coroutine can than be called with the command *StartCoroutine(name)* and cancelled at any time with *Stop Coroutine(name)*. The first coroutine was responsible for playing and looping one animation through the whole runtime of an avatar-scene, the other to play in and overlap the permanent animation with short animations. These both methods work the same way. When called, first the fitting sound was played and the matching text file was loaded. Then we took the delta time between the first to animation lines out of the text file, to receive the time the coroutine had to wait before its next run. Like that we guaranteed the same speed of the animation as we recorded it. Then each value from the text line was assigned to its matching blend shape, neck/head and eye value for the figure in the current scene. The blend shapes and eye values were assigned with no further problems, however the neck/head rotations were not that easy. Given in Quaternions we stored the start rotations in Euler-angles and then took the difference between the start rotation and current rotation. We than added the next value to the current rotation, multiplied with a scaling factor and than added the start rotation to it.

Sadly due to the fact, the text file first need to be loaded, we need to delay the audio playback. We figured a delay time of 0.2 seconds through tests, as it then was synchronized with the animation.

While we wanted the idle animation to run the whole time and been overlapped by other animations, we needed to figure a way where we did not generate jerks and get a blurred transition. Therefore we multiplied each value of the blend shapes and neck/head rotations with a scaling factor. For both coroutines, when the animation starts, the factor had the value 0 and tweened to 1 over 0.5 seconds. When an animation ended - 0.5 seconds before the sound file ended - the factor tweended from 1 to 0 over 0.5 seconds. For tweening we used the *iTween Scripting API*[34], which provided us a smooth ascent/descent of the scaling value, independent of our coroutines.

As we had 61 different animation files, we had 14 various categories of animations. Additionally we had recorded two idle animations, which were looped through the whole avatar-scene. The animations of the 14 categories than could be played in and overlapped the idle animation. The playback of these animations was done when hitting one of the keys *J, N, 1, 2, 3, 4, 5, 7, 8, 9, 0, R, F, G,* each for one category. by pressing one of these buttons, the idle animation was scaled-out, the sound stopped and simultaneously a animation of the selected category was chosen randomly. Then the coroutine for normal animations was called with the fitting text and audio file.

(a) Genesis 3 Standard **(b)** Bot Genesis

(c) Ruby Standards **(d)** Toon Generation

Figure 3.3: The different Avatars used in the experiment

3.3 Stimuli

3.3.1 Conditions

We are investigating the human behaviour when interacting with animated avatars. In detail we want to expose the changing of human behaviour, when interacting with avatars of different levels from the uncanny valley. To achieve reliable results we had to minimize factors that led to uncertain effects. Thus, we set the experiment with only two variable conditions. The first condition were four different faces seen in Fig. 3.3, including a human-like avatar (Fig. 3.3a), a robot (Fig. 3.3b), a manga figure (Fig. 3.3c) and a toon figure (Fig. 3.3d). We did not assign these figures onto the UV graph, instead we let the subjects decide it. The second condition was the motion of avatars. We presented inanimated images of the mentioned avatars and an animated versions, for the later conducted interaction.

3.3.2 Figures

All figures and blend shapes were generated using DAZ3D - a free character and posture creating system for PC/MAC. All figures were obtained from the official DAZ3D Store[35]. We decided to use DAZ3D models due to their easy model setup and their complete rigging and morphing system. Furthermore, DAZ3D includes precisely all animation blend shapes as used in the recording module of the Brekel Software. 3D Models and Renderings of DAZ3D character models were also used in previous work for stimuli creation [2].

For the 3D model of the robot we obtained the female "Bot Genesis" model. The toon character was created using the female version of the "Toon Generation" series. A more human-like character named "Ruby Standard" was chosen from the store due to their ambiguous appearance between human and toon. The human character model based on the new "Genesis 3 Standard" Model from DAZ3D. The human model contains a sub-surface-scattering map which was used for skin shading in Unity. All models consists of 4K textures (diffuse, specular, normal bump) and a complete skeleton rig. Animation blend shapes were created using the morphing system in DAZ3D. To reject lower parts of the body as well as to rename the blend shapes according to the Brekel interface in Unity all models were exported from DAZ3D to Autodesk 3ds max.

Due to performance and graphic quality issues, we baked the ambient occlusion maps onto the textures in 3ds max using the standard renderer. After that, we imported the model into Unity and rebuilt the surface shading trees. All models receive shader from the free skin and shader system "FastFakeSkin" by "PanDishPan" from the Unity Asset Store [36]. Eye-lashes, hair, and cloth meshes were shaded using the "Citonia" Double Sided Shader including specular and bump maps.

3.4 Study Design

3.4.1 Wizard-Of-Oz Room

We already mentioned the experiment design inspired by Jakubi A. Złotovski[22] in Section 2.3: The concept of the “Wizard-of-Oz Room”[37]. This concept has many advantages, including to make testing of non-existent technology possible, in our case we can simulate a futuristic AI (Artificial Intelligence) with a human-like appearance. Furthermore a fast iteration of successive runs, provides a insight into the probands interactions with the system and as the wizard you can learn about how people interact with systems. But this concept has disadvantages too, for example: The wizard need to be trained, such that he can respond in a understandable way and computers respond in a different way, than humans, even more, the wizard is able to improvise. The last fact can affect the results, in worst case.

The setup of a *Wizard-Of-Oz Room* is shown in Fig. 3.4b. The user is isolated from outer influences for the whole duration of the experiment (Fig. 3.4a), such that he can interact with the system without disturbances. Even the questionnaire he has to fulfil is

(a) Interacting User **(b)** Wizard-Of-Oz Room

Figure 3.4: Proband interacting with an avatar (left) and the Wizard-of-Oz Room Setup (right)

completed in isolation. The Wizard sits in another room or elsewhere, but is isolated from the user, and controls the system the user is interacting with.

3.4.2 Hardware

The application for the experiment ran on an Hyrican Military Gaming 4895 PC, with a intel i7-6700 Quadcore processor, with a clock speed of 3.4 GHz each kernel, 16 GB Ram and Windows 10 as operating system. With the NVIDIA GeForce GTX 970 graphic card, we ensured a smooth running system without lacks, such that the subjects did not experience bad feelings because of poor running system. The user had a mouse and keyboard to fulfil the questionnaire and enter their data. The Screen was a 24" LCD Monitor with a HD 1920 x 1080 pixels resolution. The wizard saw the same things as the screens were duplicated and connected to the same PC.

3.4.3 Interaction

We wanted to make subjects interact with a CG avatar. This interaction should be designed in such a way, that it was easy reconstructible and without much degree of freedom, but enough to get the sensation of a free conversation for the subject. Therefore we took a party game, called “WHO AM I?”. For this game, a participant slips into a fictive or real character, but without knowing which one. His quest is now to find out who he is, only with questions answerable with “yes” or “no”. This game had to be done four times, for each animated avatar. So we needed four characters for each participant. We had a total amount of 20 characters and assigned them via a Blockplan[38] to the subjects’ four avatar interactions, to make the occurrence of each character random. All

Interaction	Key	Amount	Interaction	Key	Amount
Yes	J	5	No	N	5
Yes hesitated	1	3	No hesitated	5	5
Yes won	2	3	No loss	7	5
Yes corrected to No	3	4	No corrected to Yes	8	5
Right	4	3	Nope	9	2
Hello	A	3	False question	F	2
Goodbye	Y	4	False question plus	G	2
Please repeat	R	6	No idea	0	4

Table 3.1: Answer Categories

of the assignments and characters are shown in Table A.1. Whereas the number of the figure did not indicated the actual figure, but rather the orders. To assign the avatars, with which each participant interacted with, in random order, we established another Blockplan whose construction is shown in Table A.2.

The avatar then could answer on the questions through the wizard with 14 different possibilities, matching the 14 categories we earlier alluded in Section 3.2.2. Therefore we had to record the data for the interaction. Thus we invited an actress with some experience in acting, singing and speaking. We than recorded, spread over three sessions, as already said in Section 3.2.1, a total amount of 128 different interaction answers. The different types of answers, buttons the wizard had to press to replay them later and the amount of different records in each category are shown in Table 3.1. The actress was German, so we recorded all responses in German language and held the interactions later completely in German too. So if the proband asked a question the wizard could reply by pressing one of the buttons.

3.5 Procedure

Starting a session a proband was pleased to enter and sit at the table we prepared for the experiment. We then obliged them to fulfil a consent form, where they were informed about the experiment and had to sign whether they allow us to use this data freely, anonymous or not. After that they were introduced into the Program, starting with record of the personal data: Name, Age, Profession, Sex. When hitting the Start button the proband could start the experiment by himself. Then he sees the first image of one of the four figures, inanimate, for 20 seconds, to let him be affected by the avatar. After that the wizard switches to the questionnaire, such that the proband now can rate the avatar. When finishing the questionnaire the scene switched and the idle animation for the first shown inanimated figure began. Next, the wizard pressed the “A”-button, such that the figure welcomed the proband. Than the proband started asking questions and the wizard reacted with the correct answer to them. At the end, when the proband figured the searched character or after a long time, he did not, the wizard pressed the “Y”-button, that the avatar said goodbye. These steps, without recording the personal data again, were repeated further three times, once for each avatar. At the end the proband was allowed to leave without further instructions. The complete procedure took about 25 minutes the shortest and 55 minutes the longest.

4 Results

Participants were recruited through a doodle-link[39] where one could register for participation in the experiment. Students were reached over a mailing list. On the final day, we had a total amount of 18 students, 15 male and 3 female. The youngest proband was 18 years old, the oldest 29 years, the most were between 20 and 25 years. All subjects were students, mostly of the *Univeristy of Stuttgart*.

The measurement we used, were the 19 alternative Godpseed idices' established by Chin-Chang Ho and Karl F. MacDorman[25]. The 19 indices were categorized into three parts: *human likeness, eeriness* and *attractiveness*. A quick overview of these indices is given in Table 4.1. Given eight conditions, four figures, animated and inanimated each, we received eight outputs for each proband.

A two-way between subjects ANOVA (Table A.6, Table A.7 and Table A.8) was conducted to compare the effect of human likeness, eeriness and attractiveness on virtual 3D-figures conditions. It revealed a significant effect of human likeness between the conditions [F(3, 17) = 14.544, p < .001]. As well there was a significant effect of eeriness [F(3, 17) = 6.348, p = .001] and attractiveness [F(3, 17) = 7.434, p < .001]. Furthermore the ANOVA revealed significant effects of human likeness [F(3, 17) = 24.507, p < .001] and attractiveness [F(3, 17) = 13.238, p = .002], but no significant effect on motion, due to eeriness [F(3, 17) = 3.971, p = .063]. There was no interaction effect for figure and motion for human-likeness [F(3,17) = 1.291, p = .287], eeriness [F(3,17) = .249, p = .861], and attractiveness [F(3,17) = .519, p = .671].

Post hoc comparisons (Table A.9, Table A.10 and Table A.11) using Bonferroi correction indicated that there was a significant difference of human likeness between Figure 1 (Fig. 3.3a) and Figure 2 (Fig. 3.3b) (M = .258, SD = .063) and between Figure 1 and Figure 4 (Fig. 3.3d) (M = .199, SD = .037). Also a significant difference of human likeness were indicated between Figure 3 (Fig. 3.3c) and Figure 2 (M = .186, SD = .044), and between Figure 3 and Figure 4 (M = .127, SD = .023). However the human likeness (M = .072, SD = .041) did not significantly differ from Figure 1 and 3. The same holds for the human likeness between Figure 4 and 2, no significant difference (M = .058, SD = .044).

Further the comparisons indicates a significant difference between Figure 2 and Figure 1 in terms of eeriness (M = .079, SD = .029). Similar indications hold for the comparison

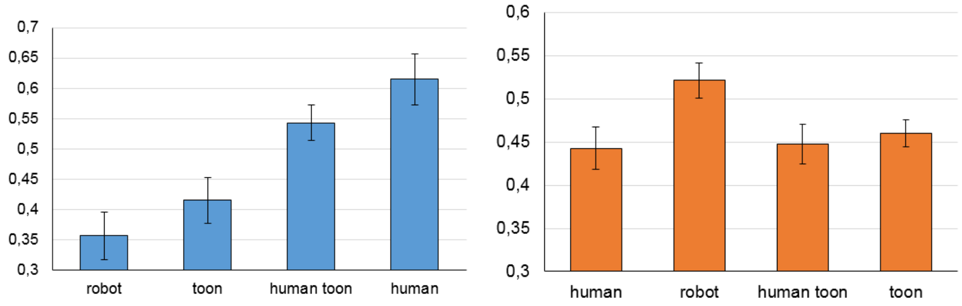
Indices			Categorie
Artificial	-	Neutral	Human Likeness
Synthetic	-	Real	
Inanimate	-	Living	
Human-made	-	Human like	
Mechanical Movement	-	Biological Movement	
Without Definite Lifespan	-	Mortal	
Reassuring	-	Eerie	Eeriness
Numbing	-	Freaky	
Ordinary	-	Supernatural	
Uninspiring	-	Spine-tingling	
Boring	-	Shocking	
Predictable	-	Thrilling	
Bland	-	Uncanny	
Unemotional	-	Hair-raising	
Unattractive	-	Attractive	Attractiveness
Ugly	-	Beautiful	
Repulsive	-	Agreeable	
Crude	-	Stylish	
Messy	-	Sleek	

Table 4.1: Godspeed indices and categories

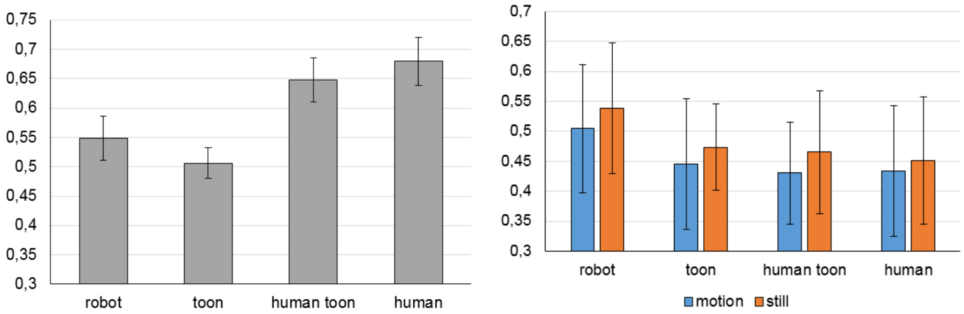
of eeriness for Figure 2 and Figure 3 (M = .074, SD = .021). Whereas the eeriness between Figure 3 and Figure 1 (M = .005, SD = .018), Figure 4 and Figure 1 (M = .017, SD = .017) and between Figure 4 and Figure 3 (M = .012, SD = .017) did not significantly differ among each other.

In regard of attractiveness the comparison of Figure 1 and Figure 2 (M = .131, SD = .047) revealed no significant difference. Likewise did the attractiveness between Figure 1 and Figure 3 (M = .031, SD = .045) not indicate a significant distinction. Equally results were established for the comparison of Figure 2 and Figure 4 (M = .043, SD = .043), and Figure 3 and Figure 2 (M = .099, SD = .040) holding the fact of no significant difference.

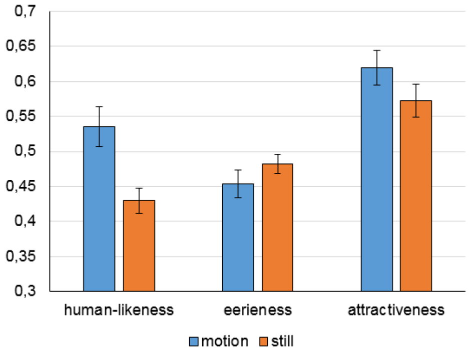
When assigning to the UV graph (Fig. 1.1), the probands rated the human likeness in such a way, that the evaluation indicates the locations of the avatars on the graph like Fig. 4.1a and Fig. 4.1b shows. The robot as the least human like ($\approx$ .36), most eerie ($\approx$0.525) on the right side of the graph and the human as most human like ($\approx$ .61) and least eerie ($\approx$ .445), thereby on the summit or ascent before the UV. Some



(a) Human likeness **(b)** Eeriness



(c) Attractiveness **(d)** Eeriness (inverted)



(e) Effect of interactivity

Figure 4.1: Visualized results of the Godspeed indices

deviations for the toon and manga figure, do not indicated clear results. Whereas the toon figure is a tiny amount more eerie (≈ 0.46) than the manga figure (≈ 0.45), that one is way more attractive (≈ 0.65) and human like (≈ 0.54) rated than the toon (≈ 0.55 in attractiveness, ≈ 0.41 in human likeness).

Fig. 4.1e shows an increase of human likeness and attractiveness of virtual figures due to interaction with them. Without regard on standard deviation, 4.1d indicates a decrease of eeriness, if human interact with animated figures.

During the experiment we also collected personal opinions and suggestions from the probands in text form. An interesting thing is, we received the least feedback for the robot figure, both animated and inanmiated. The most complains came for the human like and toon figure. Proband 5 and 13 complained about weird eye movements of the human avatar and that the figure did not looked directly at them. Subjects 1, 2 and 9 said similar things about the manga figure. The second most complains were about mouth movements during a reply, we obtained these for all figures except the manga figure. Another interesting fact is, we did not receive any feedback in relation to the interaction itself or about experiences they made, just about visual appearances.

5 Conclusion

The research question of this thesis was, if and how interactions change the way we perceive virtual 3D-Avatars. This question can be answered now, when taking the results together. We observed a significant raise of human likeness and attractiveness due to the interaction, only eeriness did not faint in such a way we wanted it, but it did a tiny bit. Hence we can say interactions affect our perception of CG avatars in an at least partly way, such that the figures affect us positively. But the fact, eeriness did not faint as we expected, just the human likeness an attractiveness raised, refutes the results of Laura M. Flach and Vanderson Dills[16], that motion only affect eeriness. Certainly we know Laura M. Flach and Vanderson Dills did not construct an interaction between CG-avatars and human, but only observing implementation, since participants of their study did not interact but observed CG avatars in motion. We could interpret that as further prove that real interaction does change our perception positive and no interaction affect it negative. Thus the submitted results of Lukas Piwek, Lawrie S. McKay and Frank E. Pollick[19] and James Thompson[20], who said there is no effect of motion, would support our thesis too.

But we cannot overrate our results due to several limitations our experiment did provide. First of all we only had 18 different subjects and only three of them female. This does not give us assurance for the reliability of the experiment. Furthermore the avatar motions we recorded, were not 100 % accurate as we wanted them to be. Another main limitation was the game subjects played with the avatar for interaction, which maybe affected the first impression. Other interactions could lead to other results. For the game, we provided some characters who should have been guessed by each participant during the interaction, but not everyone did figured them right and thus could have possible rated the interaction in a negative way. Further we recorded the amount of questions who were questioned and the time a subject needed to provide one, but we could not evaluate these data in time and we did not see benefits for us. Also limiting was the appearance of the figures as for example participant 11 reported that figure 3Fig. 3.3c had some similarities to Angelina Jolie, one of our characters and thus Angelina Jolie was easier figured than by other participants. Additionally all of our figures were female, what could have affected the outcome in a significant way, due to the fact most participants were male, hence attractiveness and sympathy were increased because of that and for the same reason eeriness did not decrease. A small limitation also was, probands wanted to be successful in the quiz and hence needed a high amount of time. Therefore the procedure took a long time. Duo to the duration and the repeated same procedure, participants were maybe fatigue and hence the ratings of the figures were not totally accurate.

So what we can take from here for future work is, create an interaction experiment without that much influencing factors and maybe shorter. Moreover create interactions with female and male avatars, to wipe out that gender could affect the interaction. We know the interaction does affect behaviour but future work could testify different types of interactions. For instance, what happens if the CG avatars insults the subject? How will unsocial behaviour of virtual avatars change human behaviour? Do we consider them as real enemies due to our primary instincts, or are we getting used to interact with CG-avatars and treat them as such unreal entities, such that we do not let ourself disturbing? Another pretty interesting thing would be, as our figures were settled on the left side of the uncanny valley, what happens to our behaviour when taking avatars out of the valley an interact with them? The spectrum of possibilities and research questions is broad at this point.

A Tables

Character Name	Character ID	Participant ID	Figures 1.	2.	3.	4.
Angela Merkel	A	1	H	N	R	E
Barack Obama	B	2	L	F	H	M
Sandra Bullock	C	3	C	K	J	R
Donald Trump	D	4	N	Q	P	A
Leonardo DiCaprio	E	5	S	I	T	O
Mariah Carey	F	6	D	B	G	E
Madonna	G	7	J	A	E	Q
Miley Cyrus	H	8	D	K	O	R
Justin Timberlake	I	9	B	P	C	S
Snoop Dogg	J	10	G	L	I	H
Angelina Jolie	K	11	M	N	F	T
Vladimir Putin	L	12	N	S	G	M
Hillary Clinton	M	13	C	P	T	B
Christiano Ronaldo	N	14	H	D	A	R
Lionel Messi	O	15	Q	E	K	F
Justin Bieber	P	16	J	O	L	I
Taylor Swift	Q	17	I	O	K	L
Jessica Alba	R	18	F	C	B	R
Will Smith	S					
Brad Pitt	T					

Table A.1: Blockplan for characters and assignments to participants

Participant ID	1.	2.	3.	4.	Participant ID	1.	2.	3.	4.
1.	A	C	B	D	10.	C	B	D	A
2.	D	A	C	B	11.	B	A	D	C
3.	C	B	A	D	12.	B	C	D	A
4.	D	A	B	C	13.	A	C	D	B
5.	A	D	B	C	14.	A	D	C	B
6.	D	C	A	B	15.	B	D	C	A
7.	A	B	D	C	16.	C	D	A	B
8.	C	D	A	B	17.	B	C	A	D
9.	C	A	D	B	18.	D	C	A	B

Table A.2: Blockplan Avatars - A = Genesis 3 Standard, B = Bot Genesis, C = Ruby Standard, D = Toon Generation

Measure		Mean	Std. Error	Interval	
				Lower Bound	Upper Bound
Human likeness	1	.615	.042	.527	.703
	2	.357	.039	.275	.439
	3	.543	.029	.482	.604
	4	.416	.038	.336	.495
Eeriness	1	.443	.024	.391	.494
	2	.521	.020	.479	.564
	3	.448	.023	.400	.495
	4	.460	.016	.427	.493
Attractiveness	1	.680	.041	.592	.767
	2	.549	.038	.469	.629
	3	.648	.038	.569	.728
	4	.506	.026	.451	.561

Table A.3: Estimates

				Interval	
Measure		Mean	Std. Error	Lower Bound	Upper Bound
Human likeness	1	.536	.028	.476	.596
	2	.430	.028	.372	.488
Eeriness	1	.454	.020	.411	.496
	2	.482	.016	.448	.516
Attractiveness	1	.619	.024	.569	.669
	2	.572	.028	.512	.632

Table A.4: Estimates

	Mean	Std. Deviation	N
AmH	.6695602	.16910479	18
AsH	.5604552	.22772378	18
BmH	.4244020	.14630365	18
BsH	.2900270	.21572947	18
CmH	.5680749	.15422041	18
CsH	.5181134	.12748474	18
DmH	.4804977	.16593868	18
DsH	.3509259	.17627506	18
AmU	.4340712	.10893628	18
AsU	.4512587	.10659285	18
BmU	.5046875	.08537088	18
BsU	.5382668	.10895206	18
CmU	.4301939	.10529128	18
CsU	.4650752	.10269285	18
DmU	.4459635	.09036615	18
DsU	.4735243	.07162675	18
AmA	.7028241	.15439269	18
AsA	.6562037	.21871771	18
BmA	.5787963	.14766468	18
BsA	.5190509	.18679091	18
CmA	.6599537	.17793121	18
CsA	.6364120	.15473597	18
DmA	.5358565	.10059692	18
DsA	.4767824	.12626127	18

Table A.5: Descriptive Statistics

Source			Type III Sum of Squares	df	Mean Square	F	Sig.
Figur	Human Likeness	Sphericity Assumed	1.490	3	.497	14.544	.000
		Greenhouse-Geisser	1.490	1.866	.798	14.544	.000
		Huynh-Feldt	1.490	2.087	.714	14.544	.000
		Lower-bound	1.490	1.000	1.490	14.544	.001
	Eerieness	Sphericity Assumed	.143	3	.048	6.348	.001
		Greenhouse-Geisser	.143	2.427	.059	6.348	.002
		Huynh-Feldt	.143	2.860	.050	6.348	.001
		Lower-bound	.143	1.000	.143	6.348	.022
	Attractiveness	Sphericity Assumed	.718	3	.239	7.434	.000
		Greenhouse-Geisser	.718	2.823	.255	7.434	.000
		Huynh-Feldt	.718	3.000	.239	7.434	.000
		Lower-bound	.718	1.000	.718	7.434	.014
Error (Figur)	HumanLikeness	Sphericity Assumed	1.741	51	.034		
		Greenhouse-Geisser	1.741	31.721	.055		
		Huynh-Feldt	1.741	35.481	.049		
		Lower-bound	1.741	17.000	.102		
	Eerieness	Sphericity Assumed	.384	51	.008		
		Greenhouse-Geisser	.384	41.256	.009		
		Huynh-Feldt	.384	48.624	.008		
		Lower-bound	.384	17.000	.023		
	Attractiveness	Sphericity Assumed	1.643	51	.032		
		Greenhouse-Geisser	1.643	47.987	.034		
		Huynh-Feldt	1.643	51.000	.032		
		Lower-bound	1.643	17.000	.097		

Table A.6: ANOVA: Univariate Tests - Figure

Source			Type III Sum of Squares	df	Mean Square	F	Sig.
Motion	HumanLikeness	Sphericity Assumed	.403	1	.403	24.507	.000
		Greenhouse-Geisser	.403	1.000	.403	24.507	.000
		Huynh-Feldt	.403	1.000	.403	24.507	.000
		Lower-bound	.403	1.000	.403	24.507	.000
	Eerieness	Sphericity Assumed	.029	1	.029	3.971	.063
		Greenhouse-Geisser	.029	1.000	.029	3.971	.063
		Huynh-Feldt	.029	1.000	.029	3.971	.063
		Lower-bound	.029	1.000	.029	3.971	.063
	Attractiveness	Sphericity Assumed	.080	1	.080	13.238	.002
		Greenhouse-Geisser	.080	1.000	.080	13.238	.002
		Huynh-Feldt	.080	1.000	.080	13.238	.002
		Lower-bound	.080	1.000	.080	13.238	.002
Error (Motion)	HumanLikeness	Sphericity Assumed	.279	17	.016		
		Greenhouse-Geisser	.279	17.000	.016		
		Huynh-Feldt	.279	17.000	.016		
		Lower-bound	.279	17.000	.016		
	Eerieness	Sphericity Assumed	.123	17	.007		
		Greenhouse-Geisser	.123	17.000	.007		
		Huynh-Feldt	.123	17.000	.007		
		Lower-bound	.123	17.000	.007		
	Attractiveness	Sphericity Assumed	.103	17	.006		
		Greenhouse-Geisser	.103	17.000	.006		
		Huynh-Feldt	.103	17.000	.006		
		Lower-bound	.103	17.000	.006		

Table A.7: ANOVA: Univariate Tests - Motion

Source			Type III Sum of Squares	df	Mean Square	F	Sig.
Figur * Motion	HumanLikeness	Sphericity Assumed	.041	3	.014	1.291	.287
		Greenhouse-Geisser	.041	2.814	.014	1.291	.288
		Huynh-Feldt	.041	3.000	.014	1.291	.287
		Lower-bound	.041	1.000	.041	1.291	.272
	Eerieness	Sphericity Assumed	.002	3	.001	.249	.861
		Greenhouse-Geisser	.002	2.513	.001	.249	.828
		Huynh-Feldt	.002	2.985	.001	.249	.860
		Lower-bound	.002	1.000	.002	.249	.624
	Attractiveness	Sphericity Assumed	.008	3	.003	.519	.671
		Greenhouse-Geisser	.008	2.228	.003	.519	.619
		Huynh-Feldt	.008	2.580	.003	.519	.644
		Lower-bound	.008	1.000	.008	.519	.481
Error (Figur*Motion)	HumanLikeness	Sphericity Assumed	.534	51	.010		
		Greenhouse-Geisser	.534	47.841	.011		
		Huynh-Feldt	.534	51.000	.010		
		Lower-bound	.534	17.000	.031		
	Eerieness	Sphericity Assumed	.120	51	.002		
		Greenhouse-Geisser	.120	42.729	.003		
		Huynh-Feldt	.120	50.746	.002		
		Lower-bound	.120	17.000	.007		
	Attractiveness	Sphericity Assumed	.253	51	.005		
		Greenhouse-Geisser	.253	37.879	.007		
		Huynh-Feldt	.253	43.856	.006		
		Lower-bound	.253	17.000	.015		

Table A.8: ANOVA: Univariate Tests - Figure * Motion

Measure			Mean difference (I-J)	Std. Error	$Sig.^b$	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
HumanLikeness	1	2	.258*	.063	.004	.071	.444
		3	.072	.041	.594	.051	195
		4	.199*	.037	.000	.088	.311
	2	1	-.258*	.063	.004	-.444	-.071
		3	-.186*	.044	.003	-.316	-.055
		4	-.058	.044	1,000	-.190	.073
	3	1	-.072	.041	.594	-.195	.051
		2	.186*	.044	.003	.055	.316
		4	.127*	.023	.000	.059	.196
	4	1	-.199*	.037	.000	-.311	-.088
		2	.058	.044	1.000	-.073	.190
		3	-.127*	.023	.000	-.196	-.059

Table A.9: Pairwise Comparison - Human Likeness

* The mean difference is significant at the .05 level.
b Adjustment for multiple comparisons: Bonferroni.

Measure			Mean difference (I-J)	Std. Error	$Sig.^b$	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
Eerieness	1	2	-.079*	.026	.046	-.157	-.001
		3	-.005	.018	1.000	-.060	.050
		4	-.017	.017	1.000	-.069	.035
	2	1	.079*	.026	.046	.001	.157
		3	.074*	.021	.015	.011	.136
		4	.062	.022	.070	-.003	.127
	3	1	.005	.018	1.000	-.050	.060
		2	-.074*	.021	.015	-.136	-.011
		4	-.012	.017	1.000	-.062	.038
	4	1	.017	.017	1.000	-.035	.069
		2	-.062	.022	.070	-.127	.003
		3	.012	.017	1.000	-.038	062

Table A.10: Pairwise Comparison - Eeriness

* The mean difference is significant at the .05 level.
b Adjustment for multiple comparisons: Bonferroni.

						95% Confidence Interval for Difference	
Measure			Mean difference (I-J)	Std. Error	$Sig.^b$	Lower Bound	Upper Bound
Attractiveness	1	2	.131	.047	.073	-.008	.269
		3	.031	.045	1.000	-.103	.166
		4	.173*	.043	.005	.045	.301
	2	1	-.131	.047	.073	-.269	.008
		3	-.099	.040	.145	-.219	.020
		4	.043	.043	1.000	-.086	.171
	3	1	-.031	.045	1.000	-.166	.103
		2	.099	.040	.145	-.020	.219
		4	.142*	.035	.005	.037	.247
	4	1	-.173*	.043	.005	-.301	-.045
		2	-.043	.043	1.000	-.171	.086
		3	-.142*	.035	.005	-.247	-.037

Table A.11: Pairwise Comparison - Attractiveness

* The mean difference is significant at the .05 level.

b Adjustment for multiple comparisons: Bonferroni.

Bibliography

[1] M. Mori, “The uncanny valley,” *Energy*, 7th ser., no. 4, pp. 33–35, 1970 (cit. on pp. 11, 12).

[2] K. F. MacDorman, R. D. Green, C.-C. Ho, and C. T. Koch, “Too real for comfort? uncanny responses to computer generated faces,” *Computers in Human Behavior*, Computers in Human Behavior, no. 25, Jan. 29, 2009 (cit. on pp. 12, 13, 31).

[3] M. Jabbi, J. Bastiaansen, and C. Keysers, “A common anterior insula representation of disgust observation, experience and imagination shows divergent functional connectivity pathways,” *PLoS ONE*, 2008 (cit. on p. 12).

[4] A. Pinar Saygin, T. Chaminade, H. Ishiguro, J. Driver, and C. Frith, “The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions,” 2011 (cit. on p. 13).

[5] R. L. Goldstone and A. T. Hendrickson, “Categorical perception,” Dec. 23, 2009 (cit. on p. 13).

[6] J. Bruckheimer. (2007). Pirates of the caribbean - at worlds end (cit. on p. 14).

[7] P. Jackson. (2003). The lord of the rings (cit. on p. 14).

[8] R. Zemeckis. (2004). The polarexpress (cit. on p. 14).

[9] ——, (2011). Mars needs moms (cit. on p. 14).

[10] (Apr. 30, 2016). Boxofficemojo, [Online]. Available: `http://www.boxofficemojo.com/` (cit. on pp. 14, 15).

[11] (Apr. 28, 2016). Spongebob, [Online]. Available: `https://wall.alphacoders.com/big.php?i=233959` (cit. on p. 15).

[12] (Apr. 28, 2016). Uncanny girl, [Online]. Available: `http://www.cubo.cc/creepygirl/` (cit. on p. 15).

[13] F. Cafaro, L. Lyons, J. Roberts, and J. Radinsky, “The uncanny valley of embodied interaction design,” 2014 (cit. on p. 17).

[14] T. Burleigh, “Does the uncanny valley exist? an empirical test of relationship between eeriness and the human likeness of digitally created faces.,” 2009 (cit. on p. 17).

[15] E. Dr. Schneider, Y. Wang, and S. Yang, “Exploring the uncanny valley with japanese video game,” 2007 (cit. on p. 18).

[16] L. M. Flach, R. H. de Moura, S. R. Musse, V. Dill, M. S. Pinho, and C. Lykawka, “Evaluation of the uncanny valley in cg characters,” 2012 (cit. on pp. 18, 19, 39).

[17] R. McDonnell and M. Breidt, “Face reality: Investigating the uncanny valley for virtual faces,” 2010 (cit. on p. 18).

[18] B. A. Urgen, A. X. Li, C. Berka, M. Kutas, H. Ishiguro, and A. P. Saygin, “Predictive coding and the uncanny valley hypothesis: Evidence from electrical brain activity,” Mar. 22, 2015 (cit. on p. 19).

[19] L. Piweka, L. S. McKayb, and F. E. Pollick, “Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion,” 2013 (cit. on pp. 19, 39).

[20] J. Thompson, G. Trafton, M. McCurry, and E. Francis, “Perceptions of an animated figure as a function of movement naturalness: No sign of the uncanny valley,” 2008 (cit. on pp. 20, 39).

[21] C. Becker-Asano, K. Ogawa, S. Nishio, and H. Ishiguro, “Exploring the uncanny valley with geminoid hi-1 in a real-world application,” 2010 (cit. on pp. 20, 21).

[22] J. A. Zlotowski, H. Sumioka, S. Nishio, D. F. Glas, C. Bartneck, and H. Ishiguro, “Persistence of the uncanny valley: The influence of repeated interactions and a robot’s attitude on its perception,” Jul. 30, 2015 (cit. on pp. 20, 21, 31).

[23] C.-C. Ho, K. F. MacDorman, and Z. A. Dwi Pramono, “Human emotion and the uncanny valley: A glm, mds, and isomap analysis of robot video ratings,” 2008 (cit. on pp. 21, 22).

[24] C. Bartneck, D. Kulic, E. Croft, and S. Zoghbi, “Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots,” 2009 (cit. on p. 22).

[25] C.-C. Ho and K. F. MacDorman, “Revisiting the uncanny valley theorie: Developing and validating an alternative to the godsped indices,” *Computers in Human Behaviour*, no. 26, 2010 (cit. on pp. 22, 35).

[26] S. A. Steckenfunger and A. A. Ghazanfar, “Monkey visual behavior falls into the uncanny valley,” Oct. 27, 2009 (cit. on pp. 23, 24).

[27] *Optitrack system*, May 8, 2016. [Online]. Available: `https://www.optitrack.com/` (cit. on p. 25).

[28] *Kinect sensor image*, May 9, 2016. [Online]. Available: `https://en.wikipedia.org/wiki/Kinect#/media/File:Xbox-One-Kinect.jpg` (cit. on p. 26).

[29] *Tobii eyex image*, May 9, 2016. [Online]. Available: http://www.tobii.com/imagevault/publishedmedia/d5vu2w7w00xexxvkeoeq/Web_Images_EyeX_2660x1280.jpg (cit. on p. 26).

[30] *Zoom h6 audio-recorder image*, May 9, 2016. [Online]. Available: https://ksr-ugc.imgix.net/assets/005/673/279/74b193844479fdaf1b7eb57370cf5d2d_original.jpg?w=680&fit=max&v=1460059841&auto=format&q=92&s=b6cebacc6af4981d690cd229ab1832cc (cit. on p. 26).

[31] (May 3, 2016). Brekel pro face, [Online]. Available: http://brekel.com/brekel-pro-face-2/ (cit. on p. 26).

[32] *Tobii eyex developers guide unity*, Mar. 15, 2016 (cit. on p. 27).

[33] May 5, 2016. [Online]. Available: http://docs.unity3d.com/ (cit. on p. 28).

[34] May 5, 2016. [Online]. Available: http://itween.pixelplacement.com/index.php (cit. on p. 29).

[35] (May 5, 2016). Daz3d, [Online]. Available: http://www.daz3d.com/ (cit. on p. 30).

[36] (May 5, 2016). Fastfakeskin shader system, [Online]. Available: https://www.assetstore.unity3d.com/en/#!/content/42535 (cit. on p. 31).

[37] *Wizard of oz*, May 6, 2016. [Online]. Available: http://www.usabilitybok.org/wizard-of-oz (cit. on p. 31).

[38] *Blockplan*, May 4, 2016. [Online]. Available: http://www.randomization.com/ (cit. on p. 32).

[39] *Doodle*, May 1, 2016. [Online]. Available: http://www.doodle.com/ (cit. on p. 35).

bibliography

All links were last followed on Mai 9, 2016.

Declaration

I hereby declare that the work presented in this thesis is entirely my own and that I did not use any other sources and references than the listed ones. I have marked all direct or indirect statements from other sources contained therein as quotations. Neither this work nor significant parts of it were part of another examination procedure. I have not published this work in whole or in part before. The electronic copy is consistent with all submitted copies.

place, date, signature