Universität Stuttgart

Implications of the Uncanny Valley of Avatars and Virtual Characters for Human-Computer Interaction

Von der Fakultät für Informatik, Elektrotechnik und Informationstechnik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

Valentin Schwind

aus Stuttgart

Hauptberichter:	Prof. Dr. Niels Henze
Mitberichter:	Prof. Dr. Bernd Eberhardt
Mitberichter:	Prof. Dr. Elisabeth André
Tag der mündlichen Prüfung:	5. Juli 2018

Institut für Visualisierung und Interaktive Systeme der Universität Stuttgart

2018

Zusammenfassung

Technologische Fortschritte machen es möglich, künstliche Figuren immer realistischer aussehen zu lassen. Oft werden sie nach dem menschlichen Vorbild entworfen, damit Interaktionen mit Computern oder Robotern natürlich und vertraut wirken. Im Jahr 1970 beobachtete der japanische Robotiker Masahiro Mori allerdings, dass Roboter oder Prothesen ab einem bestimmten Grad an Menschenähnlichkeit unheimliche, unbehagliche und sogar abstoßende Gefühle hervorrufen können. Während echte Menschen oder abstrakte Abbildungen des Menschen keine negativen Gefühle wecken, fallen Figuren, die beinahe realistisch sind in das "Uncanny Valley", wie Mori das Phänomen nannte. Mit Weiterentwicklungen in der Computergrafik sind heute auch virtuelle Figuren vom Uncanny Valley betroffen und Gegenstand zahlreicher Forschungsdisziplinen geworden. Für die Forschung sind Computergrafiken besonderes interessant, denn können einfach untersucht werden und tragen so maßgeblich zum besseren Verständnis des Uncanny Valleys und der menschlichen Wahrnehmung bei. Für Entwickler und Designer von virtuellen Charakteren ist es aber auch wichtig zu verstehen, wie die Erscheinung und der Realismus von Figuren die Erfahrung und Interaktion des Benutzers beeinflussen und wie man trotz des Uncanny-Valley-Effekts glaubwürdige und akzeptierte Figuren gestalten kann. Diese Arbeit untersucht diese Aspekte und stellt den nächsten Schritt in der Erforschung des Uncanny Valleys dar.

Diese Dissertation enthält die Ergebnisse aus neun Studien, welche die Auswirkungen des Uncanny Valleys auf die menschliche Wahrnehmung untersuchen, wie es die Interaktion beeinflusst, welche kognitive Prozesse ablaufen und welche Ursachen dafür verantwortlich sein können. Darüber hinaus untersucht diese Arbeit Methoden zur Vermeidung von unangenehmen Erfahrungen durch das Uncanny Valley und die bevorzugten Eigenschaften von virtuellen Avataren. Damit wird das Uncanny Valley in Zusammenhang mit weiteren Faktoren gebracht, die ähnliche Erfahrungen auslösen und ebenfalls vom Realismus einer Figur abhängig sind. Die Unheimlichkeit virtueller Tiere wird erforscht, wodurch Hinweise darauf gefunden wurden, dass das Uncanny Valley nicht nur auf die Dimension der Menschenähnlichkeit beschränkt ist und fundamental unsere Sicht auf das Phänomen und die menschliche Wahrnehmung verändert. Des Weiteren wurde mithilfe Hand-Tracking- und Virtual-Reality-Technologien festgestellt, welche Faktoren noch mit der Wahrnehmung von Realismus zusammenhängen. Die Affinität zum virtuellen Ich und das Gefühl der Präsenz in der virtuellen Welt werden beispielsweise auch durch das Geschlecht oder abweichende Körperstrukturen beeinflusst. Indem die Produktivität beim Arbeiten auf Tastaturen in der virtuellen Realität betrachtet wurde, ist außerdem festgestellt worden, dass die Wahrnehmung des Avatars mit den individuellen Fähigkeiten des Benutzers zusammenhängt. Diese Thesis schließt mit Implikationen ab, die nicht nur das Wissen über virtuelle Charaktere, Avatare und das Uncanny Valley erweitern, sondern nennt auch neue Design-Richtlinien für Entwickler und Gestalter von virtuellen Charakteren für die Mensch-Computer Interaktion und virtuelle Realität.

Abstract

Technological innovations made it possible to create more and more realistic figures. Such figures are often created according to human appearance and behavior allowing interaction with artificial systems in a natural and familiar way. In 1970, the Japanese roboticist Masahiro Mori observed, however, that robots and prostheses with a very – but not perfect – human-like appearance can elicit eerie, uncomfortable, and even repulsive feelings. While real people or stylized figures do not seem to evoke such negative feelings, human depictions with only minor imperfections fall into the "uncanny valley," as Mori put it. Today, further innovations in computer graphics led virtual characters into the uncanny valley. Thus, they have been subject of a number of disciplines. For research, virtual characters created by computer graphics are particularly interesting as they are easy to manipulate and, thus, can significantly contribute to a better understanding of the uncanny valley and human perception. For designers and developers of virtual characters such as in animated movies or games, it is important to understand how the appearance and human-likeness or virtual realism influence the experience and interaction of the user and how they can create believable and acceptable avatars and virtual characters despite the uncanny valley. This work investigates these aspects and is the next step in the exploration of the uncanny valley.

This dissertation presents the results of nine studies examining the effects of the uncanny valley on human perception, how it affects interaction with computing systems, which cognitive processes are involved, and which causes may be responsible for the phenomenon. Furthermore, we examine not only methods for avoiding uncanny or unpleasant effects but also the preferred characteristics of virtual faces. We bring the uncanny valley into context with related phenomena causing similar effects. By exploring the eeriness of virtual animals, we found evidence that the uncanny valley is not only related to the dimension of human-likeness, which significantly change our view on the phenomenon. Furthermore, using advanced hand tracking and virtual reality technologies, we discovered that avatar realism is connected to other factors, which are related to the uncanny valley and depend on avatar realism. Affinity with the virtual ego and the feeling of presence in the virtual world were also affected by gender and deviating body structures such as a reduced number of fingers. Considering the performance while typing on keyboards in virtual reality, we also found that the perception of the own avatar depends on the user's individual task proficiencies. This thesis concludes with implications that not only extends existing knowledge about virtual characters, avatars and the uncanny valley but also provide new design guidelines for human-computer interaction and virtual reality.

Acknowledgements

This work would never have been possible without the help and support of a number of great people. If revising my weird study designs, providing feedback for strange papers, testing crazy devices or strange setups, cleaning up the world behind me, or just talking about the boring cleaning process of the coffee machine (or my office) – without the people, colleagues, friends, and family who always get the best out of me, this work would never have been possible. I sincerely apologize to everybody I forgot to mention.

Foremost, I thank my unflinching and patient supervisor Niels Henze for showing me the amazing world of academia and scientific research. I thank him for providing unbelievable support, for his steady mentoring, and, of course, for sharing the office with me. As his first doctoral student and likely not the last one, I'm thankful for his great advise, solving easy and difficult problems, and making the world a little bit better. I also want to thank Katrin Wolf, my former colleague and the most humorous post-doc ever. Her creativity and unusual approaches gave me inspiration and offered new perspectives. Of course, I thank Albrecht Schmidt who took me into his group after my admission to the doctoral program *Digital Media*. I want to thank him for all his great support, motivating spirit, and all his recommendations, tasks, and advice I (sometimes) ignored. Of course, I would also like to thank all those who believed in me when I apllied for the doctoral program *Digital Media* and later in the SFB-TRR 161 research project:

Thomas Ertl, Daniel Weiskopf, Bernd Eberhardt, as well as my former diploma thesis supervisors Walter Kriha, Jens-Uwe Hahn, and my former collegues Beate Schlitter and Reinhold Fragner who still supported me long after studying.

"Standing on the shoulders of giants" means to get the professional and active support of many incredible people, whose work inspires and astonishes: I thank Catrin Misselhorn, Karl F. MacDorman, Christopher Bartneck, Angela Tinwell, James Hillis, Max Di Luca, Lorraine Lin, Sophie Jörg, and Marc Ernst for their great input, advise, and help. All their positive and inspiring advice are in no way inferior to the quality and importance of their work. I would also like to thank those people who have been an inspiration with their imaginative philosophy of work by improving the life of others: Rul Gunzenhäuser and Rainer Mallebrein.

Of course, I thank the crazy interaction lab/hcilab in Stuttgart crew who brightened my daily working life: Sven Mayer, Huy Viet Le, Rufat Rzayev, Alexandra Voit, and Dominik Weber. I also thank rest of the hcilab troop who is still in Stuttgart, went to Munich, or is now somewhere else: Lars "Kaffeebär" Lischke, Pawel Wozniak, Jakob "C04" Karolus, Francisco "Pepe" Kiss, Mariam Hassib, Pascal Knierim, Mauro Avila, Passant ElAgroudy, Romina "Kettner" Poguntke, Oliver Korn, Thomas Kosch, Norman Pohl, Patrick "baderp" Bader, Markus Funk, Tilman Dingler, Stefan Schneegass, Thomas Kubitza, Bastian "Basti" Pfleging, Alireza "Ali" Sahami, Tonja Machulla, Lewis Chuang, Yomna Abdelrahman, Nora Broy, Miriam Greis, Marianne Richter, Rainer Dietz, Anja Mebus, and Murielle Naud-Barthelmeß.

I am lucky to work with a number of amazing and talented students: Katharina Leicht, Nico Haas, Philipp Göttlich, Florian Nieuwenhuizen, Cagri Tasci, Patrick Franczak, Kai Friedrich, Alexandre Comeau-Vermeersch, Brighten Jelke, and Robin Schweigert.

I would like to thank my friends, who supported me more than they know: Tvrtko Grubesa, Marcus Lengerer, Benjamin Brösicke, Markus Raunitschke, Thomas Braun, and Stefan Fohrer. I want to thank my uncle, who died last year. I thank my cute but lazy cats Mitsou and Smilla, who could be cuddled at any time. I want to thank Wolf-Dieter Jäger and Monika Böttcher as parents in law and Swaantje Böttcher as the sister in law for their great support, help, and cakes. I want thank my parents Heidemarie and Rudolf Schwind, for their unlimited patience, prudence, support, and for their (dis)belief that somebody could ever get an academic degree for the crazy stuff I did in my life. Finally, to adequately conclude this, I want to thank my brilliant, supportive, and outrageously loving wife – Solveigh. She is the reason why all this happened. With her foresight, wisdom, and sanity, she made all the things possible – more than anyone could ever imagine.

Table of Contents

1

Introduction

Creating depictions of human beings has accompanied and shaped human history. Technological progress made it possible to create increasingly realistic figures. One effect that can emerge when figures look almost but not perfectly real has been coined the uncanny valley effect eliciting eerie, uncomfortable, and even repulsive feelings. As depictions of humans and other beings play a central role in many interactive computing systems, it is important to investigate the uncanny valley in the context of human-computer interaction. This chapter provides an overview of the implications of the uncanny valley for human-computer interaction – the topics of this thesis. The introduction starts with the development and importance of human depictions in human cultures. We address why today's virtual depictions of humans can increase our knowledge of perception. After defining and discriminating terminologies, we explore the uncanny valley in context of avatars and virtual characters. We summarize the contribution of this work in human-computer interaction, provide an outline of the dissertation, and the list publications used.

Figure 1.1: 3D scan of the sculpture "Warrior of Hirschlanden". The sculpture is the oldest-known human-size anthropomorphic statue with a human face north of the European Alps.

1.1 Depictions of Humans and Animals

Humans have developed the remarkable ability to depict. Unlike any other species on earth, humans use depictions of the own species for wording thoughts and conveying knowledge. The first human figurative representations are known from the Stone Age and can be considered as the beginnings of human culture. The 100,000-year-old ochre-processing workshop in the Blombos Cave in South Africa is considered as a milestone in human cognition and technology combining chemical substances for depicting and decoration [115]. The oldest preserved figurative human-like depictions can be found in the El Castillo cave (dated about 40,000 B.C.) in Spain [320] from the basal Aurignacian. Also dated back to this period is the mammoth-ivory sculpture "Venus vom Hohlenfels" found in a cave in the Swabian Jura of southwestern Germany [53]. The female figurine is the oldest-known sculpture of a human being. Another ivory sculpture from the Aurignacian culture is the "Lion-man," which was found in the Hohenstein-Stadel cave, Germany, and is the oldest-known example of figurative art. As indicated by its name, it is an anthropomorphic sculpture with a lion-head on a male human body. One remarkable example of enhanced sculpting handcraft from the Hallstatt culture is the life-sized stone sculpture "Warrior of Hirschlanden" (Figure 1.1) dated to 600 B.C., which is the oldest known human-size anthropomorphic statue that has been found north of the European Alps [247]. The sculpture is modeled with realistic features and proportions, however, it has only a schematic face which can also be considered to be a mask. These examples show that human or anthropomorphic figures can be found in the earliest stages of human civilization. Unique in nature and an integral part of human culture, the ability to create human depictions has been further improved with the technological progress of mankind.

Technological progress grew with the development of human cultures and led to increasingly realistic depictions of humans and animals. This can be seen in filigree drawings, sculptures, reliefs by the Ancient Egyptians (3100 B.C. – 332 A.D.), detailed depictions, and the Sphinx creature with its lion body and human head. An outstanding example of antique handicraft is the huge terracotta army in the Mausoleum of the First Qin Emperor of China (210 B.C.), where every soldier and horse was individually designed so that none of the 8,000 figures look identical in pose, facial features, or equipment details. Another milestone in the history of realistic depictions are human-like machines by the Arabic inventor Al-Jazarī (around 1200) [118]. His water-powered musical robot band and automated servants are among the first known animated depictions of human beings. The development of animated machines reached a spectacular climax with Vaucanson's Canard Digérateur (Digesting Duck from 1738), which consisted of over 400 moving parts and simulated a complete metabolism by eating and defecating kernels of grain [340]. The next step in the development of human-like depictions was the interactive chess-playing "Mechanical Turk" by Wolfgang von Kempelen (1770), later operated by Johann Mälzel (1804). This automaton was a mechanical hoax with a skilled chess-playing operator hidden inside the desk of the machine [302].

Two hundred years after the development of the Mechanical Turk and with the advent of the industrial revolution, artificial humans and animals became interactive and more realistic. Micro-controllers, motors, and materials were improved. Thus, engineers were able to develop very sophisticated animated human- or animal-like robots. While animal robots were used as toys, such as Spin Master's Beagle or Aibo by Sony, human-like robots were initially developed for prostheses and surgery. A remarkable example is the Collins handmade by D. W. Collins in 1961, which could be used by amputees for holding and grasping objects [240]. Today, some of the most advanced robots are the full body human-like Actroids by the Japanese roboticist Hiroshi Ishiguro. Such elaborately constructed lifelike robots are mainly developed for research purposes and to simulate credible human behavior [189].

The 20th century brought not only advances in robotics; further kinds of life-like and animated depictions emerged. The development of cameras and methods for optical recording and reproducing motion pictures brought photographs and movies. Humans and animals were now recorded and portrayed on pictures, projectors, and displays. Videos and movies that conveyed complex narrative contexts were developed. Motion pictures also enabled artists to draw manually human and animal characters from which animated films and cartoons emerged. Famous examples are the drawings by Walt Disney, who shaped the world of animated movies and the style of animated humans or animals as no one else [307]. The following digitization and rendering technologies made it possible to integrate computer-generated imagery (CGI) into movies; firstly used for post-production and special effects. Since the 90s, CGI has been used to bring virtual humans, animals, and other creatures as actors on screen. Examples are liquid metal renderings used for the almost indestructible human-like robot T-1000 in James Cameron's *Terminator 2* (1991) [36].

At the end of the 20th century, powerful and affordable hardware of personal computers (PCs) and advanced computer graphics made it possible to render virtual worlds on everyone's computer screen. However, computers not only allow rendering of realistic worlds; they also allow interaction. Using input devices and sensors, the next generation of human and animal depictions was able to either be controlled by the user or to react to the users' action. Today, we can talk with virtual characters, they can interact with us and we can explore the virtual worlds with them. In the field of human-computer interaction (HCI), virtual characters that represent users in the virtual world are called *avatars*. The function of avatars goes even beyond simply depicting humans or animals. Avatars can be created or customized by the users and help the express their personality. With the advent of the Internet, people even learned to communicate using their avatars, which wide implications in social behavior and also opened up new methods for interaction and communication [269].

In the second decade of the 21st century, advanced head-mounted virtual reality displays, body tracking technologies, and tactile feedback for everyone's use emerged. These technologies enable users to view the virtual world from the perspective of their own avatar and to render the user's body according to their real body's pose and movements. Instead of experiencing the world in their real body, people can have any appearance they like and immerse in any virtual environment they want. Thus, virtual depictions change our state of consciousness by diminishing the own presence in the physical world and create a completely new experience in the virtual one, which has far-reaching consequences for human perception. Furthermore, using virtual characters, people can slip into new roles, which has also immense effects on their personalities. Humans do no longer passively spectate or control virtual depictions – humans turn into them.

For developers and designers of interactive and immersive computing systems or applications, it is not only important to provide human-like virtual characters but also to understand how their appearance and realism influences the experience of the user. However, modified or alternative forms may evoke conflicts in our perception. Through such perceptual conflicts, researchers learn how various cognitive mechanisms in the human brain process internal as well as external cues. On the one hand, this knowledge can be used to further improve computing systems for more realistic animated life-like depictions, complex interactions, and immersive experiences. On the other hand, researchers learn more about human perception and how we can improve depictions and depicting systems.

Virtual or computer-generated (CG) characters are the most recent form of figurative human or human-like depictions. They act as a user interface between the real and virtual worlds and affect the user's interaction and perception. As they are not physical entities, they can be rendered on any display and everyone can interact with them using a computing system in real-time. However, designers and developers are faced with new challenges when creating very photo-realistic human or animal depictions or to create renderings of virtual entities, which can be no longer be distinguished from real ones.

In this work, we explore the consequences of using very realistic virtual characters and avatars in human-computer interaction. The focus of this work is on realism and human-likeness because they are the most relevant factors for credible and immersive experiences aiming to copy reality using high visual fidelity. High levels of realism are responsible for people's familiarity with the virtual experience as they are plausible and make it easier for users to feel immersed. However, and as previously indicated, designers or developers of virtual applications have unlimited freedom to vary the appearance of the virtual world and their characters. They do not necessarily invest considerable expenditures and take the trouble to make virtual worlds as realistic as possible. On the contrary, designers also can deliberately simplify, stylize, or abstract virtual characters as Walt Disney once did in his animated motion pictures. Then, the question arises which appearance represents the best trade-off between realism and acceptance of the virtual character. To further explore avatars and virtual characters, it is necessary to distinguish both and to define terms such as virtual realism or human-likeness.

1.1.1 Avatars and Virtual Characters

There is a distinction between an avatar and a virtual character. An avatar is a visual representation of the user in a virtual environment. The representation can be a 2D picture, a symbol, an icon, or a 3D model, such as a player character in video games. The avatar can be used in Internet forums or online communities and in virtual worlds such as video games or in virtual reality (VR). Avatar is the Hindu word for the "descent" of divinity in terrestrial shapes. There are discussions about the first-time use of the word avatar in the virtual context. The earliest use in a computer game was in the role-playing game (RPG) *Avatar* released in 1979. The author Norman Spinrad published his novel *Songs from the Stars* in 1980 and used the term avatar as a description of a computer generated virtual experience. Richard Garriott also coined the word in the computer game

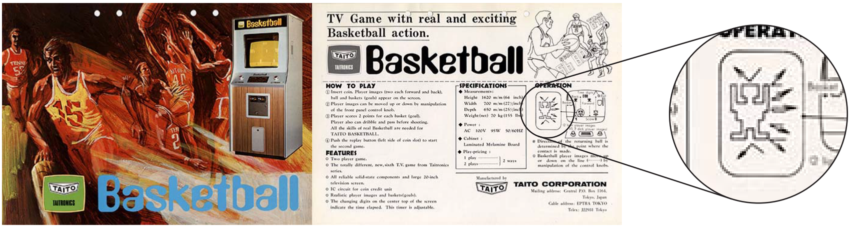

Figure 1.2: Manual of the arcade video game *Basketball* by Taito Corp. The game has the first known avatar in a video game. The human-like icon of the game is depicted on the right page of the arcade game manual. Image by Taito/Dphower. Image source: `http://flyers.arcade-museum.com`.

Ultima IV: Quest of the Avatar in 1985. Neal Stephenson popularized the use of the term avatar as a representation in an online virtual-reality application in his cyberpunk novel *Snow Crash* from 1992. However, video games used avatars before the term avatar was brought into context as virtual representations of the player. Similarly to the first use of the term avatar, there are discussions about which video game showed the *first* avatar. At the time of this work, the earliest known human-like virtual representation of a video game player and avatar was shown in the arcade game *Basketball* published by *Taito* in 1974 (scan of the manual is shown in Figure 1.2).

Virtual characters are CG figures, which represent a person or other being with human or animal characteristics. By definition, avatars are also virtual characters, but they are a special case of a virtual character as they represent only the user (*e.g.* the player character in a computer game). In interactive simulations, virtual characters are controlled by an artificial intelligence (AI). This includes non-player characters (NPCs), virtual assistants, embodied agents, and animated or simulated life-like beings with human or animal attributes. We often consider avatars and virtual characters separately, as virtual characters can represent something other than the user and avatars explicitly specify and emphasize the meaning of a virtual depiction of one's self in a virtual environment.

Furthermore, avatars do not only represent users in the virtual world; they can also be individually customized in contrast to pre-defined virtual characters by the developers or designers. While virtual characters are displayed as provided by the application, some applications allow customizations and individual configurations of an avatar by the user. Multiple customizations make the appearance of an avatar unique and allow users to express themselves. In this thesis, we often summarize avatars and virtual characters as *virtual figures*. Depictions that can also be real such as robots, dolls, and toys are summarized as *artificial figures*.

1.1.2 Virtual Realism and Anthropomorphism

As previously mentioned, increasing computing power allows to render more and more realistic graphics of virtual worlds and characters. However, when virtual humans become realistic, people do not always speak about increased *realism*, but often also about increased *anthropomorphism*. Anthropomorphic can be a synonym for *human-like* or *humanoid*. Highly realistic as well as highly anthropomorphic depictions aim to deliver a credible representation of a human being. Nevertheless, both of these terms do not necessarily mean the same.

Anthropomorphism describes the attribution of human characteristics to non-human entities. In art and animations, anthropomorphic entities receive human characteristics such as speech, emotions, upright walking, facial expressions, and other visual characteristics of humans such as clothes. Thus, increased realism does not necessarily mean that avatars and virtual character are rendered anthropomorphic. On the contrary, giving an animal a human voice does not make the animal more realistic, but more anthropomorphic such as in fables or fairy tales. Anthropomorphism is not only an artistic device of making figures more human but a systematic, psychological pattern to recognize familiar, human characteristics in behavior or appearance of non-human entities [132].

Realism in art and virtual systems aims to create or reproduce credible and truthful depictions. This includes visual qualities of an image such as lighting, shading, surface details, and resolution. Creating or reproducing visual realism such as in photos or films through art and virtual renderings is also known as *photorealism* and the artistic trend to use very high-resolution or exaggeratedly detailed photorealistic graphics using high contrast and high-dynamic color ranges is also known as *hyperrealism*.

1.2 The Uncanny Valley

In previous sections we learned about both similarities and the differences between human depictions, avatars, virtual characters, realism, anthropomorphism, and human-likeness. Directly related to the group of these terms is the phenomenon of the *uncanny valley*. In the following, we define what the uncanny valley actually is, and then discuss the challenges it brings and its meaning and importance in depicting and animating human-like entities.

This subchapter mainly refers to the articles by Masahiro Mori and includes direct as well as indirect citations of his original paper from 1970, however, unless stated otherwise, only refer to the English translation authorized by Mori in 2012:

M. Mori. "The Uncanny Valley." In: *Energy* 7.4 (1970), pp. 33–35

M. Mori, K. F. MacDorman, and N. Kageki. "The Uncanny Valley." In: *IEEE Robotics and Automation Magazine* 19.2 (2012), pp. 98–100. ISSN: 1070-9932. DOI: 10.1109/MRA.2012.2192811

1.2.1 Summary

Relationships are often described as functions. Simple and mostly obvious relationships are linear functions. In the case of realism or human-likeness of entities we might assume that the more similar an object is to a real human, the more it will be accepted and positively perceived. In 1970, the Japanese roboticist and professor at the Tokyo Institute of Technology Masahiro Mori wrote an essay on people's responses to robots that looked very human. He described a strange and counter-intuitive phenomenon he observed in the human affinity to robots.

Mori observed that people have neither positive nor negative feelings for functional machinery such as industrial robotic arms. For the functional appearance of such a robot, humans do not have any emotional connection: "People hardly feel any affinity for them." Furthermore, he mentioned toy robots with no mechanical appearance and more human characteristics perceived as pleasant and appealing: "Children seem to feel deeply attached to these toy robots." However, Mori observed negative responses to robots and prostheses with *very* human-like characteristics such as skin: "However, once we realize that the hand that looked

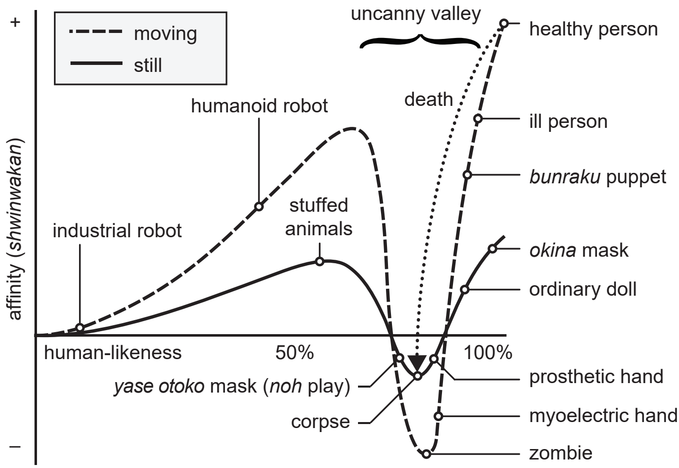

Figure 1.3: The uncanny valley. Translated and adapted version of Mori's original graph [212] from his re-publication of his essay in 2012 [213]. The graph shows the uncanny valley and the non-linear relationship between human-likeness and perceived affinity to an entity. The dotted path shows the sudden death of a healthy person. Translators' notes from the publication in 2012: *Bunraku* is a traditional Japanese form of musical puppet theater dating to the 17th century. "*Noh* is a traditional Japanese form of musical theater dating to the 14th century in which actors commonly wear masks. The *yase otoko* mask bears the face of an emaciated man and represents a ghost from hell. The *okina* mask represents an old man." [213].

real at first sight is actually artificial, we experience an eerie sensation. [...] we lose our sense of affinity, and the hand becomes uncanny." Finally, Mori mentioned Bunraku puppet play in Japan which obviously does not evoke negative feelings in the audience: "When we enjoy a puppet show in the theater, we are seated at a certain distance from the stage. The puppet's absolute size is ignored, and its total appearance, including hand and eye movements, is close to that of a human being. [...] we might feel a high level of affinity for the puppet."

Mori brought these examples together and concluded that there must be a non-linear relationship of human-likeness and affinity (Jap. 親和, *shinwakan*) and hypothesized that humanoid robots which are designed more human-like do not

Figure 1.4: The humanoid robot Repliee Q2 by T. Minato, T. Shimada, M. Ishiguro und H. Itakura realized in cooperation with the Kokoro Company. Image Credits: Max Braun from San Francisco, USA (Android), CC BY-SA 2.0, via Wikimedia Commons.

necessarily look more familiar. One example of a humanoid robot is the Repliee Q2 as shown in Figure 1.4. When such non-human entities reach a certain point of human-likeness, subtle imperfections in their appearance evoke uncomfortable feelings of eeriness or even revulsion. Even the smallest inconsistencies on face, body, or skin indicate that a person is not a human and a kind of discomfort occurs in the observer. Only when such a non-human entity is not distinguishable from a real human being does the emotional response become favorable again. Mori illustrated his hypothesis using a graph with familiarity as a function of human-likeness (see Figure 1.3). Pointing to the descent in his U-shaped graph and the evoked uncomfortable feelings, Mori called this phenomenon the *uncanny valley* (Jap. 不気味の谷現象, bukimi no tani genshō).

1.2.2 Effect of Movement

In his essay, Mori reports about a robot at the 1970 World Exposition in Osaka, Japan. One of the robots at the exposition had 29 facial “muscle” pairs (comparable to a real human) to make the robot smile. However, “[...] when the speed is cut in half in an attempt to make the robot bring up a smile more slowly, instead of looking happy, its expression turns creepy.” Regarding the uncanny valley effect, Mori hypothesized that movement of non-human entities steepens the slopes of the curve, emphasizing the positive as well as the negative affinity. Mori brought further examples such as a myoelectric prosthetic hand or moving mannequins, which could evoke horrific experiences. Without explicitly pointing to this, Mori also brought examples of interactions with humans: “For example, we could be startled during a handshake by its limp boneless grip together with its texture and coldness.” Mori also discussed death as a leap from animacy to inanimacy. The transformation of the body as indicated by the dotted path of the arrow in Figure 1.3 is visible in the pale color of the skin and the lacking of movement after we die: “We might be glad that this arrow leads down into the still valley of the corpse and not the valley animated by the living dead.”

1.2.3 Escape by Design

Mori emphasized in his essay that it is not necessary to build complex and realistic prostheses or robots if they will not be accepted. Designers risk falling into the uncanny valley when they increase the degree of human-likeness. However, by using lower degrees of human-likeness, they can avoid the uncanny valley. “In fact, I predict that it is possible to create a safe level of affinity by deliberately pursuing a non-human design.” In a broader sense, non-human design means that designers aim to simplify, abstract, or stylize the appearance of such an entity. Mori suggested that designers should aim to follow design principles that avoid high levels of human-likeness to achieve high levels of affinity. As an example for such a design principle, he mentioned a wooden statue of Buddha because “[...] it retains the natural color of the wood, but its roundness and beautiful curves do not elicit any eerie sensation.” Mori argues that artistic craftsmanship and a non-human design can prevent or circumvent the effect.

1.2.4 Potential Explanations

Coldness and pale skin tones can be considered as indicators of a serious disease and combined with lacking movement even as a sign of death. Mori observed that these aspects are somehow related to the uncanny experience, and, as previously mentioned, he referred to death as a leap between the still and moving graph (see dotted path in Figure 1.3). Mori tries to explain the phenomenon in his essay, but raises new questions: "Why were we equipped with this eerie sensation? Is it essential for human beings? I have not yet considered these questions deeply, but I have no doubt it is an integral part of our instinct for self-preservation." Furthermore, he discriminates in his essay between distal and proximal sources of dangers and assumes that "the sense of eeriness" potentially protects humans from proximal threats (*e.g.* prevention of illness from dead bodies). However, the subconscious and cognitive mechanisms as well as the underlying processes of human and object discrimination remains unknown. Mori recommends in his paper to map the valley to understand its boundaries to make the design process of non-human entities easier.

1.2.5 Challenges

Mori's original essay appeared in an "obscure" Japanese magazine and initially received no significant attention (*cf.* editor notes in [213]). In 1978, Jasia Reichardt mentioned the uncanny valley in the book *Robots: Facts, Fictions + Prediction* [240], making it known to the Western world. However, it took about three decades until the uncanny valley became a focus of scientific research. Technological advances in robotics and computer animations made it possible for humanoid robots and virtual humans to appear more and more natural. When movies such as *Final Fantasy* (2001) [261] and *Polar Express* (2004) [346] appeared and robots such as the Geminoids from the labs of Hiroshi Ishiguro were built, the uncanny valley suddenly came into public and scientific focus.

Since 2005, the uncanny valley has rapidly attracted increasing interest in robotics, animations, science, and other disciplines. Researchers have explored design implications for humanoid robots and CG humans; Biologists explored evolutionary explanations, Psychologists explored cultural and social aspects of cognition, and Neurologists searched for explanations and underlying mechanisms by investigations of the human brain. However, the uncanny valley and both dimensions “human likeness” as well as “affinity” are controversially discussed. One challenge is that both are somewhat vague concepts and difficult to quantify or operationalize. Even the translation of the Japanese word *shinwakan* seems a challenge in charting. English translations can be *familiarity*, *likeability*, *empathy*, or *affinity*, but according to an interview with the robotic researcher Karl F. MacDorman who served as the English translator for Mori’s essay, they “fail to capture the full essence of Mori’s original Japanese” [126]. MacDorman describes *shinwakan* as a “feeling of being in the presence of another human being” and the “moment when you feel in synchrony with someone other than yourself” and the “uncanny” (negative *shinwakan*) arises when this synchrony “falls apart” [126].

For empirical research, such subjective perceptions are difficult to classify. Mori’s article does not include any empirical test of his hypotheses. Today, there are studies, which confirm [187, 195], disprove [308], or correct [16, 210] aspects of his hypotheses. Current meta-research assumes that the uncanny valley is potentially a composition of multiple phenomena with overlapping cognitive mechanisms [145, 164, 268].

Today, researchers assume that the uncanny valley and its graph are not literally correct [126, 145]. Nonetheless, his hypothesis radiates certain attraction and fascination, which currently affects a considerable body of research.

1.2.6 Relevance and Impact

One important reason to understand the phenomenon and to learn its impact on human perception lies in the fact that complex human-like characters need a lot of effort to create. Even scans of real humans are difficult to capture and animate, and need an immense effort to be rendered realistically. Sometimes such characters are realized in animated movies but with considerable effort required. However, even costly animated films with very realistic characters do not perform particularly well at the box office, and the number of moviegoers significantly decreases in films with very human-like characters while stylized animations with less human-like characters celebrate enormous successes.

Independent film critics mention multiple aspects of uncanny sensations, which makes them not recommend a film [34, 50, 338]. Paul Clinton (CNN) wrote about the *The Polar Express*: “It’s a shame. ‘The Polar Express’ wants to be an uplifting holiday film, but it tries too hard to make its point. Moreover, the technology just hasn’t caught up to the lofty ambitions of the hundreds of talented people behind this film” [50]. Realistic representations and replicas of humans require significantly more effort in their creation. Related to the creation process of Princess Fiona’s character in the computer-animated movie *Shrek* (2001), Lucia Modesto (PDI/Dreamworks) made the following remark: “She was beginning to look too real, and the effect was getting distinctly unpleasant” [334]. Andy Jones, Animation Director of the first computer-animated movie with human actors *Final Fantasy* (2001), commented his work in a similar way: “As you push further and further, it begins to get grotesque. You start to feel like you’re puppeteering a corpse” [334]. Another example refers to the financial failure of the animated movie *Mars needs Moms* (2001) [333] with realistic human-like characters. With ca. $150 million budget and a box-office of ca. $39 million [28] (February 2018) the film was one of the biggest financial disasters ever. It was an immense economic disaster for Disney [182]. This is in contrast to the animated movie *Brave* (2012) [4] of the same animation studio, which appeared a few months later. With almost the same production budget and less human-like characters, the movie was with about $500 million an enormous financial success [27].

It is a fact that to the current date, there are more stylized and fewer human characters in animated films. From a scientific point of view, it is interesting to understand which factors are responsible for positive and which for mainly negative responses in CGI movies. The complexity of the phenomenon is shown in a study by Kätsyri [144], who found that effects of the critical reviews of *The Polar Express* on the perceived eeriness of filmgoers indicate that people are biased by the existence of the uncanny valley itself. Kelly Lawler from *USA Today* wrote about Carrie Fisher’s young digital double (Figure 1.5) in Gareth Edward’s *Rogue One - A Star Wars Story* (2016) [70]: “the Leia cameo [at the end of the film] is so jarring as to take the audience completely out of the film at its most emotional moment. Leia’s appearance was meant to help the film end on a hopeful note (quite literally, as “hope” is her line), but instead it ends on a

Figure 1.5: Images of Princess Leia in the first *Star Wars* Movie from 1977 (left) and from *Rogue One: A Star Wars Story* in 2016 (right). The moment at the end of the movie *Rogue One* was criticized for showing a rendered CG double. Image copyrights © 1977/2016 by Lucasfilm Ltd./Walt Disney Pictures, Image Source: Amazon Video.

weird and unsettling one" [163]. Thus, despite the financial success of the film, an emotional moment of the film was criticized due to lacking empathy with a virtual character. Similar to movies, there are also reports of "creepy" human-like robots presented at public expositions, for example [52, 165, 290]. More recently, critics of video games see the uncanny valley also occurring in titles such as *L.A. Noire* [128].

Arising questions about the acceptance of artificial figures are fundamental: What would happen if human-like virtual characters and robots appeared in our everyday life? What about assistive agents, virtual helpers, and interactive shopkeepers which are too weird? Would we accept them? Would we discriminate them like racists? For economic and social reasons, it is essential to follow these questions and to foresee answers about the uncanny valley in human-like characters. Exploring the uncanny valley also means learning more about humans. We learn more about human perception and cognitive processes which are involved in face and object perception. We learn more about human gaze behavior, aesthetic preferences, and evolutionary mechanisms causing uncanny sensations. Using VR technologies, we even learn how we perceive ourselves and how we need to improve such technologies for our future life.

1.2.7 Etymology

As previously noted, Jasia Reichardt mentioned the uncanny valley (Jap. bukimi no tani) in the book "Robots: Facts, Fictions + Prediction" [240] in 1978 where the Japanese word "bukimi" is translated into "uncanny" and "tani" into "valley" ("no" is a connecting article) [233]. In English, the usage of the word "uncanny" (beyond familiar knowledge or perception) is often used in literary and psychological contexts (in contrast to the word "eerie," which has a more general meaning of unnaturalness itself; often described as "creepy"). According to Pollick [233] and MacDorman [189], the notion of the English word "uncanny" is mainly attributed to two German articles written at the beginning of the 20th century by Ernst Anton Jentsch [139] and Sigmund Freud [86]. Pollick [233] speculates that the "uncanny" was first mentioned in *About the Psychology of the uncanny* (orig. Zur Psychologie des Unheimlichen) by Ernst Jentsch [139] who pointed out that "doubts about the animacy or non-animacy of things are responsible for an eerie feeling."

In his article, Jentsch [139] mentioned the ambiguity of automatons and their psychological effect in E.T.A. Hoffmann's pieces [123]. Sigmund Freud took up the article and criticized Jentsch, on the one hand, for not including the automaton Olimpia from the short story *The Sandman* [123] as an explicit example of *the uncanny*, and, on the other hand, because of many other uncanny motifs which are responsible for the eerie effect of the narrative (including the Sandman himself, who tore out the eyes of children). Freud himself described the "uncanny" as something once familiar, which is then hidden in the subconscious and later recurs in an alienated shape [86]. He also mentioned the uncanny "in relation to death and dead bodies, to the return of the dead, and to spirits and ghosts" [86]. Freud also linked eerie sensations with invisible manifestations and imagination. Thus, the uncanny has been associated with artificial figures and corpses long before Mori's uncanny valley. Jentsch and Freud even explained the uncanny with examples of human-like representations [86, 139]. Today (2018), the article about "uncanny" in the English Wikipedia refers to Mori's "uncanny valley" to explain the notion of this concept[1].

1.3 Motivation

Research in HCI aims to improve the interaction with computing systems and human factors. The work presented in this thesis aims to improve the interaction with virtual avatars and characters. In this dissertation, we explore how this interaction can be improved *despite* the uncanny valley. Based on the premise that avatars and virtual characters might be affected by the phenomenon of the uncanny valley and caused by realistic depictions of virtual humans, it is important to understand human perception and to distinguish the uncanny valley from other phenomena eliciting similar feelings. The uncanny valley is the leading topic of this thesis; however, as this work will show, not exclusively responsible for uncomfortable sensations when interacting with virtual characters. Thus, our research aims to improve interaction with computing systems and to consider the uncanny valley in context of HCI.

One could argue that just waiting until technical progresses enable perfectly realistic renderings and animations would avoid unpleasant effects in virtual characters. Indeed, one day there might be perfect renderings of humans that are as realistic as photos and no longer elicit any eerie effects. Foremost, it is unclear how long it would take to make such perfectly realistic graphics and indistinguishable virtual replicas. More importantly, improved technologies *per se* are not able to avoid the uncanny valley when artists have an unlimited number of design choices but no design guidelines. On the contrary, using technologies able to render highly realistic graphics further aggravates the uncanny valley when even subtle design choices of artists may significantly increase the eerie sensations. Furthermore, we must consider whether stylization can also cause discomfort in VR and if there are further factors causing eerie sensations.

The aim of this dissertation is to understand when virtual humans are perceived as unpleasant or eerie and under which conditions such characters should be used.

[1] `https://en.wikipedia.org/w/index.php?title=Uncanny&oldid=826773760`

This research is dedicated to the perception of such characters and individual differences. Main focus is if users customize their virtual characters to avoid uncanny sensations. This research aims to answer questions, which are relevant in the context of HCI and uncanny valley research and the motivation for this thesis is not only to improve interaction with computing systems but also contribute to a better understanding of human cognition. In exploring these research questions, we also come across new aspects and relevant topics expanding the understanding of human perception and human-computer interaction.

1.4 Research Questions

In a first step, we provide a general overview about the research questions (RQs), which contribute to the following topics: *explanations*, *cognition*, and *applications* within the context of the uncanny valley of avatars and virtual characters for HCI. Table 1.1 summarizes these questions according to their topic including the chapter(s) dealing with these questions. In the following, we outline why these topics are relevant in HCI and related disciplines using virtual characters and avatars.

Explanations of the uncanny valley are not only relevant to anthropology or evolutionary biology, but also in the context of HCI. An evolutionary origin of the phenomenon would implicate that probably all – or only certain – humans perceive artificial human-like entities as uncanny. Knowing that people can generally be affected by the uncanny valley, makes, for example, individual adjustments in the respective HCI context obsolete. The research question mainly asks for an explanation how, why, and when humans developed this "ability."

Cognition is necessary for humans to categorize perceived cues. Chapter 2 describes how the uncanny valley is associated with conflicts in categorization and that multiple conflicting cues seem to be involved (difficulties in categorization, inconsistencies in realism, or atypicalities). However, it remains unclear how these factors are related and if all these factors belong to the phenomenon of the uncanny valley. To better understand categorization of

Topic	No.	Research Question	Chapter
Explanations	RQ1	What is the explanation for the uncanny valley?	2, 3, 5, 6
Cognition	RQ2	Under which conditions does the uncanny valley occur?	5, 6
	RQ3	Does the uncanny valley only apply for human-like depictions?	5, 6
Applications	RQ4	Which methods can avoid potential uncanny effects?	3, 4, 5, 6
	RQ5	How does the uncanny valley of the own avatar affect immersion in VR?	4, 6
	RQ6	Does VR immersion with altered body structures depend on avatar realism?	6
	RQ7	Does VR immersion depend non-innate abilities *and* avatar realism?	6

Table 1.1: Summary of research questions of this thesis: The contributing topic, research questions, number, and the corresponding chapter(s) targeting the question.

human-like and non-human-like entities, it is fundamental to understand if the uncanny valley not only occurs in human-like but also in animal-like entities. From an HCI perspective, this questions is particularly important as an uncanny valley of animals would have a significant impact on the design of virtual animal characters or animal-like robots.

Applications are concrete use cases and mainly relevant for HCI researchers, developers, and designers. Design guidelines and tools can help to avoid or overcome the uncanny valley. However, previous work derived design principles only by deduction and associations with features causing unpleasant effects. Thus, research in HCI is interested in learning how to actively avoid uncanny effects. To answer further questions in HCI, it is fundamental to explore the uncanny valley in VR as this technology allows a direct link between the user, the appearance of the avatar, and the experience of the user. The research questions are related to the conditions under which the uncanny valley can occur in virtual reality and how it affects immersion.

1.5 Research Contributions

Through investigating the RQsshown in Table 1.1, this thesis makes three main contributions corresponding to the research topics of these questions, which are mainly but not exclusively relevant for uncanny valley research and HCI.

We contribute with findings that support *explanations* stating that the uncanny valley is likely to have an evolutionary origin (RQ1). In an eye-tracking study, we show that the uncanny valley changes gaze behavior in a similar way as shown for other primates. Human attention was drawn by the eye regions, which are the most important feature in facial recognition and categorization if an entity is real or not. Eye contact is reduced in eerie human stimuli; associated with an evolutionary mechanism to avoid threats and manifest as an appeasing behavior. We also contribute with explanations of how aesthetics and morbidity modulate affinity as outcome of the uncanny valley.

We investigate the uncanny valley to understand under which conditions it occurs. We contribute with findings that the uncanny valley not only occurs in human but also using virtual animal entities. Prior research assumes that atypical features trigger unpleasant effects only in human-like entities [145]. However, we show that decreased levels of realism and atypicalities can cause uncanny effects for non-human-like stimuli and depend on each other, which has also been shown for human-like entities (RQ2). Thus, we found support for the findings that the uncanny valley is a mechanism, which does not just depend on the dimension of human-likeness but on typicality and consistency of human as well as animal entities (RQ3). This contribution is particularly relevant for designers of applications in HCI that make use of virtual animals. We contribute with design implications and recommend to provide consistent degrees of realism and no atypical features at high degrees of realism for both human and non-human entities (RQ2). We also contribute with findings in *cognition* showing how humans incorporate human-like avatars in VR. We show that the dimension of human-likeness must be extended to the individual appearance and the user's body in VR as deviations from the own appearance can also cause eerie effects (RQ5, RQ6, RQ7). Lacking coherence between the virtual and the own appearance may cause negative affinity ratings towards the avatar and hinders immersion in the

Tool	Description	Chapter
faceMaker	Based on the European average face, people have the opportunity to design a virtual character online, in 3D, and in real time. They decide on the facial parameters of their favorite face. Researchers can give different objectives to learn which face is behind a concept. An avatar-downloader in *faceMaker* allow to save, render, and share the generated 3D models and motivates users to contribute. We used *faceMaker* to learn how people avoid uncanny effects when the have a free design choice for virtual faces.	4
Avatar Hands in VR	This prototyping tool developed in Unity3D enables researchers to create immersive experiences with low-cost hardware for hand and finger tracking using the Leap Motion sensor mounted on the HMDs. The system provides six different hands and three different tasks and can easily be extended to study immersion and avatars in VR.	6
Questionnaires in VR	This Unity module helps designers and researchers to ask participants items on 7-point Likert scales in virtual reality. Advantage of this module is that e.g. presence questionnaires can asked without leaving the VR. Thus, participants must not rely on their memory while using post-test questionnaires, which reflects an inconsistent and incomplete picture of the VR-experience. The module can be used with the virtual hands module in which subjects can fill in the questionnaire using the virtual hands whose effect should be measured.	6

Table 1.2: Developed prototyping tools in this thesis.

virtual world. Thus, we contribute to the field of cognition with findings that enhance our understanding of human-likeness and how the user incorporates the avatar into his or her own body scheme.

Finally, we provide methods to avoid the uncanny valley for virtual characters in several *applications* (RQ4). Such methods include avatar customization, aesthetics and the preferred characteristics in virtual faces. We also contribute with design implications for virtual characters and avatars. Additionally, we contribute with three prototyping tools (see Table 1.2) and provide further recommendations and design implications for first-person avatars to increase presence in VR and other immersive applications.

1.6 Methodology and Evaluation

Throughout this work, we gathered quantitative and qualitative data in user studies. Such data was collected in laboratory setups and online surveys. Based on previous work, we developed interactive and immersive prototypes with high functionality and high visual fidelity to explore the occurrence of the uncanny valley and related effects. We developed interactive online applications which worked under real conditions with hundreds of users. We selected visual fidelity of the stimuli and prototypes according to our research questions. We took objective and subjective measures using a set of established methods and tools from the literature. Statistical analysis was performed using accepted approaches from literature and previous work. For hypotheses testing, we used parametrical and non-parametrical tests at a significance level α of .05.

1.7 Publications and Work Distribution

This dissertation is based on a number of prior scientific publications:

- The literature review in Chapter 2 of the author was initiated by Catrin Misselhorn, editor of the book *Collective Agency and Cooperation in Natural and Artificial Systems* in the *Philosophical Studies Series* [272].
- The study in Chapter 3 was mainly driven by the author. The study was performed with the help of Solveigh Jäger, responsible for participant acquisition and data collection. The paper has been published in the proceedings of *Mensch und Computer (MuC)* [273]. As our work received a Best Paper Award, the article was additionally published in the *i-com* Journal by the same publisher [274].
- Chapter 4 was driven and initiated by the author. The chapter contains two studies based on the same apparatus. The apparatus was part of a student research project by Verena Dengler, Katharina Leicht, and Julia Sandrock as part of their student project at the Hochschule der Medien (HdM) Stuttgart. The apparatus was further improved by the author and has been published

on github.com[1]. Data analysis and paper writing were supported by the co-authors. Valuable input for concept and framing by Niels Henze and Katrin Wolf resulted in articles published in the *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '15)* and in the book *Game Dynamics* [277, 278].

- Both studies in Chapter 5 are based on the Masters thesis by Katharina Leicht supported by a collaborative research project with Solveigh Jäger from the HdM Stuttgart and the University of Stuttgart. The students were responsible for stimuli creation and data collection. The author initiated and drove the project. The author wrote the paper with the input and support of Niels Henze and Katrin Wolf. The paper has been published in the *International Journal of Human-Computer Studies (IJHCS)* [279].
- Chapter 6 includes three studies using similar experimental setups. Software of the first two studies are based on versions of the same apparatus which has been published on github.com[2,3]. The apparatus was developed with the help of Cagri Tasci, Patrick Franczak, and Nico Haas. Cagri Tasci helped with the additional technical development for the second study. The third study of this chapter is based on the Bachelor thesis by Florian Nieuwenhuizen. All articles were written with support of Niels Henze and Pascal Knierim [150, 281, 285]. Lewis Chuang contributed with additional input to the second study. While Anna Feit contributed with her experience in the third study, Pascal Knierim was main author of the third article and mainly drove the project while the author was in a research intern. The first of the three studies was published at the *2017 Conference on Human Factors in Computing Systems (CHI '17)* and received an Honorable Mention Award at the conference. The second study was published at the *2017 Conference on Computer-Human Interaction in Play (CHI PLAY '17)*. The third user study of this chapter will be published at the *2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*.

[1] `https://github.com/valentin-schwind/facemaker`
[2] `https://github.com/valentin-schwind/selfpresence`
[3] `https://github.com/valentin-schwind/lessfingers`

- Chapter 7 includes the conclusion and a summary of this thesis. The summary of implications and design guidelines is accepted for publication in the *ACM Interactions Magazine (Sept/Oct '18)* .
- A number of newspapers and magazines reported on findings presented in this work: SPIEGEL ONLINE (2017-08-13), WELT (2017-08-14), heise online (2017-08-12), Stuttgarter Zeitung (2017-07-20), Stuttgarter Nachrichten (2017-07-20), Badisches Tagblatt (2017-08-14), Südwest Presse (2017-08-15), Sächsische Zeitung (2017-08-17), Main Echo (2017-09-4), Schwäbische Post (2017-08-15), Schwäbisches Tagblatt/tagblatt.de (2017-08-15), Ostsee-Zeitung.de (2017-08-24), Rhein-Neckar-Zeitung (2017-08-14), svz.de (2017-09-02), Pirmasenser Zeitung (2017-09-07), Schwarzwälder Bote (2017-08-16) basic-one.com (2017-08-13), Vorarlberger Nachrichten (2017-08-26), Salzburger Nachrichten (2017-08-14), Frankfurter Neue Presse (2017-09-06), Taunus Zeitung (2017-09-06), Nassauische Neue Presse (2017-09-06), Höchster Kreisblatt (2017-09-06) , Rüsselsheimer Echo (2017-09-06) , Norddeutsche Neueste Nachrichten (2017-08-24), Oberhessische Presse (2017-08-19), Schwarzwälder Bote (2017-08-16), Neckar-Chronik.de (2017-08-15), Heilbronner Stimme (2017-08-14), Münchner Abendzeitung (2017-08-14), Neue Osnabrücker Zeitung (2017-08-14), Ruhr Nachrichten (2017-08-14), Halterner Zeitung (2017-08-14), Münsterland Zeitung (2017-08-14), Dorstener Zeitung (2017-08-14), Stuttgarter Zeitung (2017-07-21), Vorarlberger Nach- richten (2017-08-26), Badische-Zeitung (2017-08-15), Freie Presse (2017-08-29), Neue Württembergische Zeitung (2017-08-15), Hohenzollersche Zeitung (2017-08-15), Metzinger Uracher Volksblatt (2017-08-15), Reutlinger Nachrichten (2017-08-15), Alb Bote (2017-08-15), Haller Tagblatt (2017-08-15), Hohenloher Tagblatt (2017-08-15), Rundschau für den Schwäbischen Wald (2017-08-15), Allgemeine Zeitung Main (2017-08-14), Darmstädter Echo (2017-08-14), Wiesbadener Kurier (2017-08-14), Gießener Anzeige (2017-08-14), Lampertheimer Zeitung (2017-08-14), Aar-Bote (2017-08-14), Idsteiner Zeitung (2017-08-14), Gießener Anzeige (2017-08-14), Main-Spitz (2017-08-14), Wormser Zeitung (2017-08-14), Wiesbadener Tagblatt (2017-08-14), Usinger Anzeige (2017-08-14), Münsterland Zeitung (2017-08-14),

Dorstener Zeitung (2017-08-14), Neue Württembergische Zeitung (2017-06-08), Heidenheimer Zeitung (2017-06-08), Metzinger Uracher Volksblatt (2017-06-08), and Rundschau für den Schwäbischen Wald (2017-06-08).

- Further publications in scientific journals, conferences, and workshops by the author that go beyond the scope of this thesis include topics such as visual-haptic integration in VR [282] social acceptance of VR glasses [280, 284], eye-tracking [60, 176, 277], game content creation [154], touch interaction [116, 166], display technologies [10–12, 283, 339], and smart home environments [326].

1.8 Thesis Outline

This dissertation comprises seven chapters, the bibliography, and the enumerating lists. This work contain the results and evaluations of a total of seven empirical studies conducted between January 2014 and September 2017 as well as an additional literature review in the related work chapter. This work is structured as follows:

Chapter 1: Introduction contains the description of the topic, an overview about the research questions and contributions, and this outline.

Chapter 2: Background and Related Work is a summary of previous work in the field and provides insights of the cognitive frameworks and explanations. This chapters also includes a literature review providing a historical and cultural retrospective of previous work before Mori coined the term reporting about uncanny and uncomfortable sensations of human-like depictions and devices in the human past.

Chapter 3: Importance of Eyes presents an investigation of the uncanny valley using eye-tracking. The study was conducted to understand how changes in gaze behavior are related to eerie characters and how they can help to identify a potential evolutionary origin of the uncanny valley and support research aiming to explain underlying cognitive mechanisms.

Chapter 4: Preferred Virtual Faces contains two interactive studies, where participants determine the preferred characteristics of virtual faces. We show how attractive and appealing characteristics of virtual human faces can avoid uncanny effects. The chapter mainly contributes to the development of design guidelines for avatars and virtual characters in the field of HCI.

Chapter 5: Virtual Animal Characters investigates if the uncanny valley is only related to the dimension of human-likeness and significantly contributes to the field of cognition. In two studies using virtual animals, we gain more knowledge about the context of the uncanny valley, the cognitive processes, and which paradigms of the uncanny valley can be transferred from human-like to animal-like depictions.

Chapter 6: Avatars in Virtual Reality investigates the impact of the uncanny valley in HCI and explores further phenomena in VR causing similar sensations. In three subsequent studies using first-person avatars in VR, we show how the experience of being in another body and with different degrees of realism is influenced by gender, body structure, transparency, and the individual task performance.

Chapter 7: Conclusion is the final chapter contextualization the findings and contribution. We provide design implications and suggestions for further research.

2 Background and Related Work

The research presented in this thesis is rooted in uncanny valley research, computer science, as well as cognitive psychology. This chapter presents the background and findings from these disciplines. It starts with a historical and cultural literature review. We show that ambiguous, artificial figures were deliberately used by narrators, artists, and engineers to elicit eerie effects. We learn more about the cognitive frameworks explaining the uncanny valley and explanations that propose an evolutionary origin. We learn foundations of categorical perception and why the perceptual mismatch hypothesis of inconsistencies in human realism could be an explanation of the uncanny valley. We present related work by looking into the effects of movement and the impact of the uncanny valley on human interaction.

Parts of this chapter are based on the following publication:

V. Schwind. "Historical, Cultural, and Aesthetic Aspects of the Uncanny Valley." In: *Collective Agency and Cooperation in Natural and Artificial Systems*. Cham: Springer International Publishing, 2015. Chap. 5, pp. 81–107. DOI: 10.1007/978-3-319-15515-9_5

2.1 Historical and Cultural Backgrounds

Due to the fast technical progress in robotics and computer-animations, the uncanny valley has gained increased importance in the last four decades. This period of time is relatively short when we assume that the uncanny valley has an evolutionary origin and could have been observed earlier. Thus, there are authors that assume that the uncanny valley is just a temporal or generational phenomenon [31, 220] due to very realistic depictions of robots and virtual characters in recent times. They argue that the uncanny valley disappears when people get used to such depictions.

Homo sapiens did not evolve with robots or computer-animated characters. However, in search for explanations of the uncanny valley (RQ1), it is important to understand if the phenomenon is perceived differently across different generations. Answering this question is helpful to determine whether the uncanny valley appears, disappears, or even modulates over time. Thus, we ask if there were uncanny sensations with realistic human-like characters before its discovery in 1970 [212]. In the following, we explore the historical emergence and contemporary impact of the uncanny valley and understand that the phenomenon is strongly connected to historical, aesthetic, and cultural influences of the human species. The review shows that artificial figures have always been associated with unconscious fears of annihilation and the loss of the own identity – similar to the fears that are today associated with the uncanny valley [31, 192]. The emergence of such fears can be caused by an evolutionary mechanism explaining the uncanny valley. Examples from history of art, engineering, and literature demonstrate that this fear is an essential human feature and can be associated with a deep-rooted mechanism that will affect our behaviors and our decisions potentially as long as mankind exists.

In this section we consider the uncanny valley from a new perspective, which turns it from a hypothetical and marginal issue of researchers to a relevant aspect of culture and history with an impact on the sociocultural development of mankind. We assume *a priori* that Mori's hypothesis relates to human-like representations in general; not only to robots or virtual characters, but also to visual arts and written narratives that have to be envisioned in the mind's eye. By comparing these narratives and contemporary reports, we are able to understand why today's image of very human-like artificial figures is negatively affected, how social aspects increase sensitivity in perceiving artificial figures, and how the uncanny valley may have significantly shaped the human development.

2.1.1 Artificial Humans in Ancient Cultures

The Sumerian poem written in the 3rd millennium B.C. about Gilgamesh, the king of Mesopotamia, is one of the earliest extant poetry and tells the story of Enkidu, who was made of clay by the goddess Aruru to undermine the strict and extravagant reign of Gilgamesh. Enkidu is a primitive form of man who is close to nature and lives together with animals. He is discovered by a poacher, and by order of the king who already knows about him, people send Enkidu a woman. He alienates himself from nature and becomes socialized, undergoing the stages of human development (from nomads, to farmers, and shepherds), and finally meeting the king in the capital Uruk. After a fight they become friends and live through many adventures together. Gilgamesh's mother adopts Enkidu, so that the king and Enkidu become brothers.[1] After killing the Bull of Heaven (a mythical creature in several myths of ancient times), the gods punish Enkidu for his deeds with a deadly disease he later painfully succumbs to. Enkidu disintegrates into dust and Gilgamesh is in deep mourning for a long time. After Enkidu's death Gilgamesh recognizes that he is mortal, too. Thereafter, the search for eternal life becomes the leitmotif of the *Epic of Gilgamesh* [196].

Enkidu is created by craftsmanship and divine magic (similar to many other creation myths). One of the main topics in the ancient story is the character's helplessness in collision with the environment, which in the end brings him to a tragically fatal end. He becomes an instrument in the hands of gods who want to give Gilgamesh a lesson in humility. In the end, Enkidu pays with his life for the withdrawal from nature and for becoming an unscrupulous hunter. This story has no direct evidence that would speak to a special eerie effect of Enkidu's role. However, this figure causes particular fascination: He is created artificially by the gods, is an outsider, and makes humans aware of their mortality. Later, we will see that these are the characteristics typical for other artificial figures. Another

[1] In the Sumerian version of the poem, Enkidu remains a slave and servant of Gilgamesh.

interesting issue about the Gilgamesh story is that Enkidu's death is caused by his creator. For many other artificial figures later – especially those with personality and intelligence – this conflict of existence is automatically preprogrammed. In the following, will understand that this is an important aspect of the uncanny valley phenomenon.

In Greek mythology, the creation of artificial life using magic and clay plays also an important role. As the Titans Prometheus and his brother Epimetheus take a walk, they see the divine potential of the earth and so they form animals and man from clay. Every animal gets a special talent, and humans get every special quality from all the animals. The goddess Athena, a friend of Prometheus, recognizes the potential of his work and gives people sense and reason as a special gift. Prometheus is so proud of his creation that he becomes a patron and teacher to humankind. However, other gods, led by the king of the gods, Zeus, are against an emancipated species and demand sacrifices and worship from the people. Prometheus turns against the other gods and brings divine fire to the people. The punishment by Zeus follows immediately: he gives an order to his son Hephaistos, god of craftsmanship, to create an artificial woman: Pandora. She is blessed with all kinds of gifts, she is seductive and beautiful. She possesses a vessel[1] with a disastrous content that shall instigate sorrow and death over mankind. Hermes brings Pandora to Epimetheus who succumbs to her magic and opens the vessel although his brother had warned him not to accept any presents. Since that day mankind has been struck by illnesses, disasters, and sorrows again and again. Furthermore, Zeus orders Hephaistos to bind Prometheus to a rock in the Caucasus Mountains. An eagle pecks the immortal in the liver day after day. Prometheus is freed several millennia later by Hephaistos. Meanwhile, the cycle of creation and annihilation of mankind repeats several times until the children of Prometheus and Epimetheus, Pyrrha and Themis, create people from stone – "a hard race and able to work" [271].

The ancient Greek poets and philosophers Aischylos, Hesiod, and Platon, as well as the Roman Ovid provided different versions of the legend of Prometheus. The myth about the creation of the human species tells about the rebellion against

[1] The vessel is usually known as the "Pandora's box." The term has come about through a translation error. The Greek word *pithos* originally referred to a big amphora used for water, wine, oil, or grain.

the divine order and the attempt to develop a self-determined culture. This theme appears most fascinating and encouraging to many artists. Johann Wolfgang von Goethe used this theme in his famous poem *Prometheus in the time of Storm and Stress* [29, 98]. “In the Age of Enlightenment, the poem acts as a firebrand – a well-articulated contempt for all inherited or self-proclaimed authorities” [94]. As we have established earlier the idea that artificial characters would or can fight for freedom and self-determination appears uncomfortable to most of us. Perhaps the Greek gods on Olympus were adjudicated to have a similar unpleasant sensation. The motif of the rebellion of a self-determined artificial species produced in series and suitable for work was taken up again in the 20th century, when seemed technically possible to build an entire class of robots. In other times artificial figures were rare and individual products with a special status. Pandora is a remarkable example for the first negatively associated artificial figure in history: She embodies our fear of manipulation and disastrous intention, in this case covered by the seduction abilities she embodies as a woman. Despite different traditions, we still have an accurate picture of Pandora: She is seductive and equipped with many gifts such as beauty, musical talent, curiosity and exuberance. Aphrodite also gives her gracious charm, Athene adorns her with flowers and Hermes gives her a charming language [117]. Both fascination for her beauty and fear of her gift remain vivid today and represent probably the first manifestation of the strange effect elicited by an artificially created figure.

Beauty is a quality frequently mentioned in relation to artificial figures. One example is the poem by by Ovid’s *Metamorphoses* about Pygmalion, a Cypriot sculptor [3]. After having had bad experiences with sexually licentious women, he withdraws into confinement and carves an ivory statue in secret. Pygmalion falls in love with the realistic and life-sized statue and treats her as if she were a real woman. He cares for her: he dresses, adorns and fondles the figure in a loving way. At a celebration in honor of the goddess Venus (the Greek Aphrodite) Pygmalion asks in a prayer that gods give him a real wife who looks like his statue. Venus fulfills his wish; when he comes back home and kisses the statue, the ivory becomes warm, soft, and alive. The figure awakens to life and becomes a living female human.

The *Pygmalion* story is one of the most popular poems in Ovid's Metamorphoses (1 B.C. – 8 A.D.). Ovid tells about three motifs for the creation of an artificial figure: aesthetics, loneliness, and love. For the first time, material gets a special meaning: the noble and organic ivory underlines the natural and aesthetic claim of the figure. Through the centuries, the Pygmalion motif has been innumerably transferred and reinterpreted by poets, painters, and musicians. Until 1762, no text mentions the name of the statue. In the very influential work by Jean-Jacques Rousseau *Pygmalion, scène lyrique* dated in 1762, the sculptor swears eternal fidelity to the statue that was first named "Galate" which means "Milk White" [255]. Noteworthy to mention in this context is the operetta *The Beautiful Galatea* by Frank Suppé where the statue transforms from a virgin to a psychotic nymphomaniac until the goddess Venus transforms Galatea back into a statue [61].

The term *Pygmalion* or *Rosenthal effect* becomes established later in psychology as the outcome of a self-fulfilling prophecy [253]. For example, anticipated positive assessments by a teacher ("this student is highly gifted") are subconsciously transmitted and confirmed by increased attention. Furthermore, pygmalionism describes sexual affection towards human representations in the form of statues, paintings, and dolls, which can also serve as a fetish. Life-sized human replicas are currently produced commercially as sex dolls (*e.g.* "Real Dolls" made of silicone). Although these dolls are very realistic, they can be nevertheless uncomfortable to people without pygmalionism (*cf.* Valverde *et al.* [321]). The attraction of artificial figures dominates the effect of the uncanny valley [107, 108]. The intense bond can also be formed by visual and haptic contact, which is frequently the case of lonely men [124]. The habituation also could be an explanation when "[...] the stimuli continue for a long period without unfavorable results" [310]. The British TV documentary series *The Secret Of The Living Dolls* [305] shows how frightening the living dolls can be for the viewer. It shows people who live as dolls in whole body dresses made of silicone. Critics and the audience describe the documentary as extremely disturbing, creepy, and scary [25, 205, 303, 335]. Maybe Pygmalion foresaw this effect when he asked

Figure 2.1: Pygmalion and Galatea (ca. 1890). Oil on canvas by Jean-Léon Gérôme (French, Vésoul 1824–1904 Paris), The Metropolitan Museum of Art, New York, Public Domain (CC0 1.0), online available at: `http://www.metmuseum.org`.

Venus "shyly" for a real wife like the ivory virgin, and not for a living statue. Venus transformed the statue into a real human made of flesh and blood – the divine way out of the uncanny valley.

2.1.2 Artificial Humans in Non-Western Cultures

A series of books[1] dated around 350 B.C. are attributed to the Daoist philosopher Lieh Tzu. In Book V – The Questions of Tang a story is told about the automaton of the engineer Ning Schï, who is presented with his human-like figure to King Mu of Chou at his travelling court. First the king does not understand what the inventor wants to show him because he considers Ning Schï's construction to be an ordinary person. The machine can sing, dance and do various tricks. However, when the automaton makes advances to the concubines, the king cannot bear it and wants to execute him immediately. The engineer disassembles the machine to demonstrate that it was only composed "of leather, wood, glue, paint, from white, black, red and blue parts" [96]. The king orders Ning Schï to reassemble the machine, examines the mechanism and recognizes its various functions. "For a sample, he removed the heart and the mouth could not speak anymore; he removed the liver and the eyes could not see anymore; he removed the kidneys and the feet could not walk anymore" [172]. The king is extremely impressed and wonders, "how men can reach the works of the creator." The king takes the machine into his wagon and drives back to his home. At the end of the story, two masters of engineering, who thought they had already reached the limits of the humanly possible, are so impressed by the story about Ning Schï's machine that they are afraid to speak ever again about their craft as an "art."

A story in the Chinese *Tripitaka* (a collection of educational writings of Buddha) probably has the same origin as the five sons of the king Ta-tch'ouan have different talents. The first son is clever, the second son is inventive, the third one is handsome, the fourth is vigorous, and the fifth is always very lucky. The sons decide to travel to various kingdoms to find out which of their "extraordinary

[1] There are few English translations. The most well-known summaries of Lieh Tzu's stories about automats can be found in "Science and Civilisation in China: Volume 2, History of Scientific Thought" [219] and *The book of Leh-tzŭ* [96, 101]. A comprehensive translation of Lieh Tzu texts was derived from the German sinologist Richard Wilhelm. This summary refers to his original German translation from *Das wahre Buch vom quellenden Urgrund* [172].

virtues is the most outstanding" [328]. The inventive brother goes to a foreign kingdom and manufactures a mechanical man out of 360 parts (mostly wooden). He gets a lot of respect for his work and is blessed with gifts. The machine sings, dances, and acts. When the king of the land and his wife hear about that craft, they go up to a tower and have a look at the wooden man. They are both very amused but do not see a wooden man until the actor winks upward to the queen. The king orders to cut off the head of the man. Nevertheless, the inventor, who is the "father" of the figure, cries and says how much he loves his "son." He holds himself responsible for his son's poor education, begs for mercy, and finally switches off the machine. The inventor pulls a pin and the mechanical man breaks into its components. The king wonders how he could be fooled by an artificial man. He claims that the inventor has a gift "which is unequaled in the world" and gives him tons of gold. The inventor returns home and distributes the gold among his brothers. In a song he praises his work and boasts: "Who is able to surpass me?"

Both examples show that stories about the creation of artificial figures are not limited to Western culture. Simply the idea that there might have been an artificial mechanical man in China 2300 years ago is remarkable and noteworthy. The story of Ning Schï's machine and the wooden man of the son King Ta-tch'ouan are inspired by the enthusiasm that such high art of engineering evokes. The creation of an artificial human being is regarded as the greatest gift ever. King Mu was explicitly interested in functionality of the apparatus and kept the machine unceremoniously for himself. Technical scholars also show reverence for the difficulties of the building process of such complex machines and redefined the craft, which they had previously regarded as art. Today we cannot know for sure whether such machines truly existed in ancient China and whether these figures were really indistinguishable from a real human. However, if people really had never seen anything like that before, they had to assume that it was a clad or painted man – and regarded it to be an actor – like it happens in the second story.

Historical reports of apparatuses that have been used for entertainment came not only from China but also from Arabia. The fall of the Roman Empire constitutes at the same time the end of an epoch of many literary and technical achievements. In the beginning of the 9th Century, the Caliph of Baghdad initiated preservation and translation of ancient writings (*Graeco-Arabica*) and thus made an invaluable contribution to the conservation of Greek science and philosophy, which was also characterized by fascination and interest in anthropomorphic machines. In the 12th century the Arabian engineer and author Al-Jazarī continued antiquity research in his *Book of Knowledge of Ingenious Mechanical Devices* (in particular he uses the knowledge of the pressure and suction of water by Heron of Alexandria) and creates a detailed manual for the construction of such machines [1]. Many clocks, fountains, doors, locks, etc. that are preserved until today prove to be functional. Among these machines are barkeeper dolls and a machine with four mechanical figures sitting in a boat singing and playing musical instruments. Figures are mostly painted and made of mounted copper, wood and possibly of papier-mâché [119]. Some of Al-Jazarī's figures could move their heads, arms, and legs. There exist different interpretations of the effect these figures produced: on the one hand, automated mechanisms apparently delayed the movement of the figures, on the other hand "subtle caprices" of the characters resulting from these delays may well have been intentional [218]. Obviously some of the figures served for amusement and entertainment. The decorations, the intricate design manuals, as well as the high number of contemporary translations in the Arabian region indicate a high popularity of such machines and the fascination they evoked. The book by Al-Jazarī represents only one of the highlights of the epoch of technical achievements and inventions in the Arabian world.

2.1.3 Demons and Automata in the Middle Ages and Renaissance

Stories of human-like figures in the European Middle Ages are predominantly coined by the Christian and Jewish faiths. There are traditions of the golem, legends of mandrake roots with human-like forms, and alchemical instructions for creating a homunculus. Well-known are several myths about the *Prague Golem*, a mute Jewish legendary figure who grazed through the cities before the Jewish festival of Passover. A note on the Golem's forehead or in his mouth brought him to life and kept him under control. According to a legend, once Rabbi Löw forgot to remove the note so the Golem was able to walk through the city without any control. In one legend, it was possible to tear the note so that the golem crumbled into a thousand pieces [328]. The medieval reports have some things in common: artificial figures are created under mystical or demonic influence. These figures have little to do with divine creation and have less aesthetic appeal than ancient Greek statues or the elaborately painted machines of medieval Arabia. They were not used for entertainment or amusement but often had a repulsive effect. In legends the misshapen Golem does not have the ability to speak or to develop a free will. According to Hildegard von Bingen [311], the devil lives in the mandrake root, which has to be exorcised with spring water. And in most traditions the homunculus is only about the size of a fetus and could be bred on an organic substance like blood, flesh, excrement, sperm, or urine (from the epilogue by Völker [328]).

During the European Middle Ages the concept of artificial life is associated with demonic powers and negative response to the efforts of alchemists [109, 157]. According to a legend, scholar Albertus Magnus "constructed a door guard of metal, wood, wax and leather" [302]. The guard welcomed visitors with the Latin "Salve!" and asked for the reason of their visit before they were allowed to step in. One day Magnus' pupil, the young Thomas Aquinas, smashed the door guard angrily into pieces. There are different specifications about the possible reasons: some say he was so scared of the android that he had smashed it with a stick. Others say he did not want to listen any longer to the "chatter" of the guard [328]. Anyway, Magnus was very upset that Thomas had destroyed "'the work of 30 years."' Thomas Aquinas wrote years later that a soul is a prerequisite for any proper motion and demons are responsible when "necromancers make statues speak, move and do other things alike" [6].

Both demons and the undead are myths that have existed since before the Middle Ages (*e.g.* the Lamia from Greek mythology), and these myths are spread beyond Europe (*e.g.* Asanbosam from West Africa or Jiang Shi from China). However, there is no evidence that all these stories are due to a ubiquitous fear of death or to the uncanny valley (*cf.* MacDorman [183]). The first extant reports on demons in the shape of living corpses, known as vampires today, come from Southeast European countries [167]. The peculiarity of vampire stories is how the appearance of the living dead is reported: pale skin, unnatural eyes – these are a few but visible abnormalities which distinguish undead from ordinary people. This allows to establish a link between these stories and the elements that seem responsible for the experienced eeriness of objects that fall in the uncanny valley; the reasons why nowadays artificial figures assigned to the uncanny valley resemble those we find in play in stories of vampirism. The combination of visible signs of pathogenic diseases and ambigu-ity between life and death prompt negative emotions such as the eerie feeling of the uncanny valley. It is unlikely that the living dead have harmed anyone, but it is quite possible that conscious or subconscious fear of death inspired people to invent such stories, myths, or figures. Such fears can function as a warning of potential threats or socially harmful behavior in times when the mortality rate was very high. In times of prosperity and security, these fears may appear superfluous and only fascinate us, especially if we are affected by forbidden stimuli.

Between 1495 and 1497 Leonardo da Vinci presumably built a functional robot which had a complicated mechanism hidden under a knight's armor. Pulley blocks and cogwheels driven by hydropower put the robot's arms in motion [254]. Da Vinci's construction plans have survived until today and display a strong influence from the Arabian entertainment automats [302]. Greek and Arabian works on mechanical devices were revised and translated in Europe only around the end of the 15th to the beginning of the 16th century. An extensive and well-known work from Europe of this period is *Les Raisons des Forces Mouvantes* by the Frenchman Salomon de Caus. De Caus illustrated many constructional ideas; moreover, he took up plans from Greek and Arabic engineers and built elaborate machines himself. Many designs were exhibited with great success in Paris and at the Heidelberg Castle. In the palace of the Duke of Burgundy in Saint-Germain, de Caus built a system with a total of 256 artificial figures or machines driven by water power. Particularly popular motifs were wheel-driven animated scenes from Greek mythology [302]. Such machines became fashionable at courts and in the big cities of Europe. De Caus' constructions became very popular and were often copied as props for theatre performances. During the Renaissance, machines were socially acceptable and were popular elements of garden and park ensembles.

In the middle of the 16th century an Italian clockmaker and engineer Juanelo Turriano developed special virtuosity in construction of machines. He became famous for his water lifting device in Toledo. According to a legend, Turriano built an artificial figure that even went shopping with him [302]. This story was very persistent and gave the street where Turriano lived its present name: *Calle del Hombre del Palo* (the avenue of the wooden man). In his time, Turriano had to defend himself because of the accusations of an abbot who was convinced that Turriano was in league with the devil. However, the design of machines developed further without ceasing till the end of the 16th century, and machine builders were competing in the production of increasingly sophisticated and more and more spectacular and entertaining figures. Noteworthy to mention here are the automata of Vaucanson, which made their creator very rich, as well as the clockmaker family Jaquet-Droz, who were the first to develop interchangeable program rollers for their figures. The *Three Musicians* by the Jaquet-Droz family are still in good working condition and are displayed in the Museum of Neuchatel in Switzerland.

At the beginning of the 17th century, human-like machines were very common and were treated like modern day pop stars. Philosophers, doctors, and anthropologists started to be interested in such constructions. In 1637, French philosopher René Descartes predicted that people would eventually be able to develop a soul-less machine that would look and behave like an animal. He compared a mechanical pumping process with the blood circulation of animals and also drew comparisons with humans. A legend tells that Descartes built an eponymous android child as a replacement for his illegitimate daughter Francine. In 1649 when Descartes was invited to the court of Christina of Sweden he took his constructed "Francine" with him. On his journey from Amsterdam to Stockholm, the suitcase with the android drew the attention of the superstitious sailors. When the ship was caught in a storm in the North Sea, Descartes was accused of standing in league with the devil and was made responsible for the storm. The captain ordered that the android be thrown overboard [302].

As previously mentioned, a machine that attracted a lot of attention[1] was the "Chess-playing Turk" dating back to the 1760s. This mechanism, built by Hungarian Baron Wolfgang von Kempelen, may be considered to be the first

[1] E.T.A. Hoffmann's story "The Automata" took von Kempelen's Chess Turks as a model. Contemporary literary critics say that Hoffmann's fascination with the Turk does not affect the reader [95].

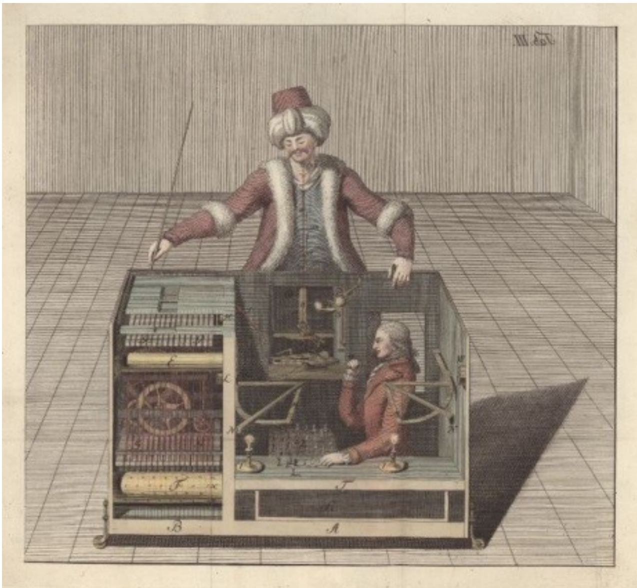

Figure 2.2: The Chess-playing Turk (1789). From *The book that tried to explain the illusions behind the Kempelen chess playing automaton (The Turk)* after making reconstructions of the device. Author: Joseph Racknitz, with friendly permission of: University Library, Humboldt University of Berlin, 3639 v.

machine of uncanny intelligence. A figure of a male android in a Turkish costume would sit in front of a box with a chess board mounted on it. The machine was designed so that when a chess move was made, the android responded with a move of a chess figures itself. Some claim the construction was able to say the word "Chess!" and "Gardez!" [302, 328]. Von Kempelen claimed that he had succeeded in developing artificial intelligence equal to the chess-playing abilities of humans. From 1783 to 1784 he traveled with his machine through Europe and let the "Chess-playing Turk" compete with well-known chess players.

Von Kempelen stopped showing the machine in 1785 after Frederick the Great offered him a large sum of money for unveiling the secret and had apparently been quite disappointed by the solution of the puzzle. After von Kempelen's death the Turk fell into oblivion for a few years until the German inventor Johann

Mälzel purchased the machine and demonstrated it again in Europe and USA as "The Automaton Chess Player." The Turk won many games during those demonstrations and defeated some famous statesmen such as Napoleon and Benjamin Franklin. At a presentation in London, the English poet and author Edgar Allen Poe observed the machine closely. In his famous essay *Maelzel's chess player* [231] he made it clear that the machine must have been a swindle and thus exposed the fraudulent automaton. Apparently both Mälzel and von Kempelen used children or midgets sitting in a small box inside the machine.

Machines in the Renaissance and Enlightenment reached a considerably higher level of acceptance than those designed in previous centuries. They were built for entertainment and, since they acted autonomously and performed tricks in front of the eyes of spectators, were a big attraction for the audience. Presumably, this fascination was particularly triggered by curiosity when the audience – as in the case of the Chess Turk – tried to find out how figures worked. However, society as well as religion did not allow crossing a certain border: an artificial figure could not be designed as one-to-one copy of a human being – and in turn a human was not allowed to be presented as a machine. That became particularly evident in the case of French physician and materialist Julien Offray de la Mettrie who was inspired by the theories of Descartes and described a man as a kind of machine for the first time in 1745 [204]. Mettrie was an atheist and polarized the world with his theories of a demystified human existence, as he not only explained complex bodily functions with mechanical processes but also saw the soul as a result of physical development. He opposed Descartes' dualism of mind and matter. La Mettrie was persecuted by the clergy and his works were heavily censored.

Enlightenment poets and philosophers such as Friedrich Schiller, Denis Diderot, and Jean-Jacques Rousseau set themselves against La Mettrie instead of giving him their support [138]. Frederick the Great took La Mettrie as his "court atheist," but Voltaire, who also worked at court, ridiculed the physician [141]. La Mettrie retired and fell into oblivion until the 19th century. The developments in this period show that automation was widespread and accepted – especially in entertainment. The controversy around the construction of artificial figures was more intense the more human-like a machine looked. The idea that a human being can be seen as a machine or can be reproduced as a machine provoked extremely strong rejections, even in times of Enlightenment and in a world with many automata. This era makes it particularly clear that both enormous technical progress and the fear of artificial humans tend to coexist.

2.1.4 Uncanny Creatures in 19th and early 20th Century Literature

The uncanny effect produced by artificial characters – particularly the idea that they could look human or threaten humans by their very existence – is often used in dark stories of the romantic period. Seen from the perspective of literary science, E.T.A. Hoffmann's *The Sandman*, published in 1815, is a "special discipline of representatives of all methodological directions" (in the epilogue by Hoffman [123]) of its own – not only in the uncanny valley research. Jentsch and Freud refer to Hofmann's piece for their explanations of eerie feelings. Nathanael, a mentally disturbed student from a good home, falls in love with Olimipia – the beautiful "daughter" of the physicist Spalanzani – and observes her through the window of the house opposite. To see her more closely, Nathanael uses a spotting scope, obtained from the glass dealer Coppelius. From the beginning of the story, Nathanael suspects the glass dealer to be the Sandman – an eye-stealing nightmare figure he knows from his childhood. Although the daughter does not move and sits motionless at the window, through the telescope Nathanael regards the doll as a living being. At a celebration Spalanzani presents Olimipia to the public and the guests realize that she is a mechanical and therefore artificial wooden doll. Although she is able to play the piano and is part of the entertainment, she looks fake with a stiff expression that makes some guests feel uncomfortable. Nathanael does not recognize the imitation because he is so much in love with the android and wants to make her a proposal of marriage. Olimpia's beauty outshines her imperfection in behavior. Spalanzani and Coppelius fight over the doll and its glass eyes fall out and scatter on the floor. The torn-out eyes remind Nathanael of the Sandman. Succumbing to madness, he attempts to kill Spalanzani. Nathanael is detained and transferred to an asylum.

As Freud [86] noted, originally Nathanael's fears are not implicitly attributed to the doll but to his fear of the Sandman. The disturbing effect on his behavior is tremendously enhanced by his love for the doll, which he sees as a real woman, especially by viewing her through the enlarging lens. In Hoffmann's story, the use of such a perspective seems to be a possible way to avoid the uncanny valley. The motif of a broken glass eye and the subconscious fear of castration due the loss of the eyesight enhance the eerie effect [86]. The real horror for Nathanael begins when he realizes that Olimpia is not a real human. Similar to Pygmalion's statue, Olimpia is a product of male fantasy, which was designed according to the wishes of a man and is the object of a man's desire. Passionate love makes both Nathanael and Pygmalion blind to the fact that their figures are not real humans. Hoffmann deliberately uses the uncanny effect produced by an artificial figure to create the atmosphere of horror and fear in his story. This method is taken up again in today's science fiction and horror stories remains a popular method to trigger deep-rooted fears. Particularly interesting in that case is Hoffmann's personal affinity for machines. Machines fascinated him: "Once when the time will be, for the benefit of all sensible people I see with me, I will make an automaton" [113]. As Nathanael's character proves, Hoffmann's highly pronounced and contemporary fascination for artificial human-like figures shows the frightening ambiguity when an "idea of an imitation of man by machine turns to a vision of horror" [214].

Today's most famous artificial figure in world literature appeared three years later in one of the first science-fiction novels ever [85, 215]. Mary Shelley's *Frankenstein, or the Modern Prometheus* written in 1818 tells the story of Victor Frankenstein, who is obsessed by the idea of the creation of an artificial human. Frankenstein works sloppily and compiles his figure out of the body parts of criminals and material from the slaughterhouse. He reanimates the body with electricity and creates a three-meter tall monster that is ugly and scary. Because of shame and fear Frankenstein keeps his creature secret, and this causes numerous serious problems. The beast sees himself as a victim and asks Frankenstein to create him a woman with whom he wants to escape from civilization. However, Frankenstein fears that together with his wife the monster could kill even more people and be a danger to future generations. Frankenstein destroys the almost completed figure of the monster's wife. The monster takes revenge by killing Frankenstein's bride Elisabeth and flees. Victor wants to hunt him down and follows the creature up to the Arctic. During his travel, Victor becomes seriously ill and dies. The creature returns to him and commits suicide when it becomes conscious of its terrible deeds [289].

Warning against too much enthusiasm and the irresponsible use of modern technology is a feature that characterizes the period of gothic novels at the beginning of the 19th century, when the topic of artificial figures was particularly important. The general eerie effect produced by the gruesome story is reinforced because the nameless monster is composed of corpses. The image of a monster is being deliberately exploited in Frankenstein not only to initiate suspense and horror but also to highlight the dangers of a human trying to take over the role of God. Such eerie stories grew popular in that period along with its widespread machine manufacturing. We can assume that highly realistic copies of human-like figures or reports about them not only triggered an eerie and disturbing effect, but also served as an indication of the potential dangers and as advice to deal carefully with the technical heritage of the antiquity. A novel aspect of Frankenstein's monster is its autonomous, uncontrolled behavior, which illustrates the powerlessness of a (human) creator towards his creation [153]. The image of Shelley's monster had a strong influence not only on the literature of dark romanticism, but also on many generations thereafter. To the 21st century it remains a highly controversial topic. Shelley's and Hoffmann's figures have been templates and role models for numerous theatre adaptations, plays, films to this day, and continue to be deliberately used to express our deep-rooted fear of human-like artificial figures and to issue warnings against excessive technological advances and irresponsible actions of science.

Between the 19th century and the early 20th century, there was a further period of particularly complex human-like machines produced by known magicians. The machines could demonstrate various magic tricks. Famous designers of that period were the founder of modern magic Jean Eugène Robert-Houdin, French magician Stèvenard, Jacques-Henri Rodolphe, Jean David, and brothers Maillardet. The figures of this period had a strong appeal not only on stage. During the period of industrialization, especially in Paris, a small industry of automatic machines emerged, so enthusiasts and collectors could purchase artificial figures. With the outbreak of World War I the industry came to an abrupt end and the era of artistically designed, complex, and human-like amusement machines was over.

After the First World War, the role and effect of artificial figures particularly depended on whether the country was a winner or a loser in the war. Victory led to euphoria and an optimistic approach to technological progress, whereas the defeated countries were generally very skeptical about it. The Czech author Karel Čapek connected this fear with the vision of a collective of artificial entities. Influenced by the subject of the Prague Golem from Jewish mysticism, he addressed the use and danger of artificial figures to warn of a further World War in his play *Rossumovi Univerzální Roboti (R.U.R.)*, whose premiere was in 1921. The play centers on a company that manufactures robots to be used as a cheap workforce. The influential utopian drama describes social and global economic consequences of the widespread usage of robots. The robots finally rebel against oppression and extinguish mankind in a terrible war. The play was a major success worldwide and was translated into almost thirty languages [152]. The remarkable aspect about Čapek's play is that artificial, human-like figures are used not for amusement but for hard work, which people are not willing to do any longer. On the other hand, machines are not artistically-crafted, individual productions anymore, but uniform entities of serial mass production. The use of robots has an enormous social and economic impact on the world, and the rebellion of the machines even leads to the end of mankind. In his play Čapek warns against the power of political concerns and dictatorships and against the power of artificial intelligence. Karel's brother Josef, a painter and writer, coined the title of the play. The Czech word "robota" stands for forced labor and at the same time is a synonym for a human-like artificial apparatus, which should ease the people's cumbersome work. The fate of the two politically engaged brothers is tragic: Karel died as a result of a hunger strike as he demonstrated against the Munich Agreement in which the Allies decided to surrender the Czech Republic to Germany. His brother Josef kept on demonstrating against the seizure of power by the Nazis through numerous performances with the play R.U.R. [306]. He was murdered in the concentration camp Bergen-Belsen in 1939.

2.1.5 Artificial Humans in Animated Movies and Films

In 1927, Fritz Lang made the ambiguity of artificially created robots the topic of his expressionist silent movie *Metropolis* [158]. In the eponymous city there are two societies: the upper class living in luxury and the working underclass in the lower parts of the city. The city is administrated by the sole ruler Joh Fredersen,

Figure 2.3: Brigitte Helm in Metropolis (1927) [158], Remastered, Public Domain (CC0 1.0), Image source available at: `http://archive.org`.

whose son Freder falls in love with Maria, a worker woman and preacher who lives in the lower part of the city. Frederson wants to suppress the rebellion at the early stage by replacing Maria with a machine-man that the scientist Rotwang has constructed. Fredersen compels Rotwang to make the robot look like Maria. However, Rotwang, who is driven by revenge, reprograms the machine so that it incites the workers to rebel against the authorities. The rebellion succeeds, and the crowd rushes in a nerve-wracking chase through the city, running into the real Maria, who tries to appease the mob. The workers accuse her of being a “witch" and drag Maria through the streets. Meanwhile, the machine-man is thrown on the pyre and the human-like shell of the machine burns. As the metal is exposed, workers recognize the fraud and that Rotwang had misused them, pursuing his own purposes. They chase Rotwang and Fredersen to the roof of a cathedral, where they start fighting each other. After Rotwang falls down, Maria mediates between the workers and Frederson and restores peace.

Brigitte Helm played the part and embodied both figures in an eerie double role; helpful and benign Maria and the sexually unbridled machine-man that is the personification of Rotwang’s sinful plans to manipulate the frustrated workers.

It was the first time that a robot embodied the uncanny doppelganger motif in a film, which was described by Freud as fear-inducing in "the highest degree" [86]. However, in America the reviews of *Metropolis* were devastating. H. G. Wells wrote in the *New York Times* on April 17, 1927: "I have recently seen the silliest film. [...] It gives in one eddying concentration almost every possible foolishness, cliché, platitude, and muddlement about mechanical progress and progress in general served up with a sauce of sentimentality that is all its own" [332]. The review was written and printed at a time when the newspapers were full of reports about the tragedy of the Titanic; the passenger liner had sunk exactly one week before and is seen today as emblematic of the hubris of people delighted with the possibilities of technology. Nevertheless, both Metropolis and R.U.R. firstly indicated that machines are able to replace humans one day due to superiority in intelligence, strength or mere number.

After World War II, computers and artificial intelligence strongly influenced the view on robots and artificial characters. It had been already been established that androids would need an extremely high intelligence and a huge amount of computing power in order to interact with their environment. For emotions, however, it was completely unclear what could be calculated by machinery. One of the first figures focusing on this distinction is the Tin Man from Lyman Frank Baum's fairy tale *The Wonderful Wizard of Oz* [19]. The Tin Man, part of the fellowship in the story about Dorothy Gale, wants a heart to be able to feel emotions. Since then, science fiction writers and film directors have been making use of the emotionlessness of artificial characters or intelligences to produce drama or weirdness. For example, in *Do Androids Dream of Electric Sheep?* by Philip K. Dick [59] a group of androids which can hardly be distinguished from humans go out of control. These androids are unable to simulate emotional reactions, which can only be determined with a complex detector – the "Voigt-Kampff machine."

The science fiction thriller *Blade Runner* (1982/1992) directed by Ridley Scott [286] is based on another of Philip K. Dick's stories and is a paragon of how to deal with sinister figures. Here, the androids are called replicants whereas one of them – Rachael – is unaware of its artificiality and appears helpless and pitying, other androids are well aware of their superior skills and use them, thus posing a threat to humans. They are chased by the Blade Runner – a bounty hunter. In a test screening the audience was unsatisfied with the end of the film, and producers insisted on changing it to a “happy ending.” The director’s cut [286], however, indicates that the Blade Runner himself is a replicant and has to flee together with Rachael. The attitude toward artificial protagonists changed quite similarly in the science fiction series *Star Trek: The Original Series* (1966-1969) and *Star Trek: The Next Generation* (1987–1994) by Gene Roddenberry [30]. Whereas in the original series, artificial intelligences or figures were mostly just eerie, highly intelligent antagonists, in the following series the emotionless android Data was a full member of the crew. The search for humanity and emotions of Data became a leitmotif of the series and the subsequent movies.

Horror films have also made use of the terrifying effect of human-like figures. Film critic Steve Rose writes: “[...] but film-makers have known about it long before it had a name. It’s what makes many horror movies tick. Zombies are archetypal monsters from the bottom of the uncanny valley, with their dead eyes and expressionless faces. Likewise the glazed-over doppelgangers in *Invasion of the Bodysnatchers* or the robotic *Stepford Wives*, not to mention the legions of dolls, dummies, puppets, waxwork figures and clowns that have struck terror in the hearts of horror fans, from the ventriloquist’s dummy in *Dead of Night to Chucky in Child’s Play*” [252].

The Scandinavian series *Real Humans* (2012) [14] shows that only a subtle adjustment is needed to make real actors look like eerie robots. The author of the series Lars Lundström explains in an interview how the special gesture play of the “Hubots” (human household robots) developed: “For this we actually needed a long time. [...] Finally we consult a mime actor to learn how to decompose movements and recompose them liquidly again. Then we thought how we could avoid all the small human gestures: no blinks of the eyes, no scratching or touching of the own face, an upright posture. Basically, the actor had to act normal, but in an abnormal way. The Hubots act like humans, but you can see that there are no real people. They are somewhat like bad actors: you can exactly see what they are trying to do” [129].

The uncanny effect of artificial figures can be specifically transmitted through the actors’ craft, but also occurs unintentionally. This can happens with obviously artificial protagonists with whom an emotional connection to the audience should be made. In this case the uncanny valley phenomenon can be held responsible if there is no emotional bond of the target audience with artificial actors. According to Misselhorn [208] movies require a kind of "imaginative perception," *e.g.*, the spectator only imagines perceiving something, but does not really perceive it. In animated characters that fall into the uncanny valley and cause a feeling of eeriness, imaginative perception gets in conflict with real perception. The uncanny valley is often mentioned as a reason why films like *Final Fantasy* (2001) [261], *The Polar Express* (2004) [346], or *The Adventures of Tintin* (2011) [295] are criticized and have not achieved box office success. The Disney production *Mars Needs Moms* (2011) [333] even counts as one of the biggest flops in film history [13]. Studies on uncanny research also use CG images to investigate the eerie effect [40, 43, 192, 297, 314].

A particular feature of artificial figures in the 20th century is that their intelligence works without figurative representations or representative bodies. Only the voice and indirect actions of HAL9000 in *2001 – A Space Odyssey*, both in the book [49], as well as in the eponymous film version [156], are sufficient to create an oppressive atmosphere. Man can hardly prevail against the uncanny intelligence in the background. Other examples of undefinable forces in the background sending deadly humanoid machines to fight mankind: Skynet from *Terminator* (1984) [37] sends a cyborg from the future into the present to wipe out the human race; similarly, the Wachowski Brothers let the last survivors of humanity fight against *The Matrix* (1999) [330] and their virtual agents. However, machines and their artificial intelligence are mostly man-made and display significant weaknesses. Many of these figures or intelligences have become an integral part of pop culture and major trademarks.

Whereas artificial figures and intelligences are seen rather negatively in the West, in Asia especially in Japan "where cultural perspectives on robots have developed rather differently from perspectives in the West" [191] exists a more positive attitude toward robots. This is possibly due to the rapid technological development of Japan, which has relied on robots since the industrialization of the country and robots do not constitute a threat to jobs there. In Japan, advanced household robots are considered as therapeutic and are being increasingly used to care for the elderly; a peaceful application of robots. Nevertheless, the Japanese affinity for robots has its limitations: 11 years after the discovery of his uncanny valley Masahiro Mori wrote: "[...] when the negative qualities of human beings are multiplied by the negative qualities of a machine, the results can be catastrophic" [211].

2.1.6 Summarizing the Historical Review

The historical review reveals that communication and cooperation between humans and anthropomorphic figures or machines have not always run smoothly. We will address the positive examples later, but in general, negative or skeptical experiences influence our view of the encounter between real and artificial humans. Artificial figures are often exemplarily used to warn of the consequences of rapid advancements in technology. Running into danger to be vanquished by an uncontrollable species often triggers an existential fear. The question is how our imagination of artificial figures has been influenced by the uncanny valley. Are artificial figures only eerie because of their negative image in history or because of their negative impact due to the uncanny valley? With hindsight, we look back at a divergent picture of the uncanny valley in history, because relying on handed down reports can hardly provide with an accurate idea of what artificial characters from the past really looked like and what people really felt when they saw them. However, in the moment of a sensory impression as well as in stories, we always try to get a concept of a figure or person in our minds. Only a few have ever seen an android, but many will have formed a negative, neutral, or even positive view on androids. However, both direct sensory perception, as well as the notion of an ambiguous shape or entity, may arouse uncomfortable feelings. Conceptual as well as imaginative perception (*cf.* Reid *et al.* [241]) of artificial figures are influenced by hitherto neglected factors: intention, aesthetics, and the cultural context of a figure.

Intention: As mentioned in the beginning, artificial figures are often designed according to our expectations and simulate human appearance or human behavior. This mental model [168] also includes a kind of awareness and intention [84, 347]. Amusement machines made by the Chinese, the Arabs, and during the Renaissance in Europe were primarily described as entertaining devices or just as tools. The lack of awareness leads to no rejection, because there is no active threat against humanity. So, these stories tell of no further conflict between man and machine; however, characters such as the men of *Prometheus*, Čapek's robots in *R.U.R.*, the androids from *Blade Runner* etc. show that self-determination of an autonomous and emancipated species is not accepted by the predominant one and treated as a serious threat that often results in a devastating conflict. Mary Shelley's *Frankenstein* deals with a figure's self-determination and describes the artificial figure as a monster with apparent cruel intentions and moral errors. This not only decreases the emotional bonding of the reader with the tragic figure, but also increases the reader's doubts and fears. Artificial figures like Pandora, Olimpia, or the Machine from *Metropolis* also produce an eerie image when distracting (often with their appearance) from an evil purpose or leaving the reader or protagonist in the dark about their true intention. This might, in principle, also apply to von Kempelen's/Mälzel's "intelligent" Chess-playing Turk, because of its ambitious intention to win the game against humans. In contrast, neutral or philanthropic intentions in combination with tragic fates like the death of Gilgamesh's companion Enkidu, the end of Magnus' doorman, Descartes' daughter, or Rachael in Blade Runner, may appear less eerie and even pitiful for human beings.

Aesthetics: Because people are accustomed to associating their counterparts to a specific gender, androgynous artificial figures are nearly always portrayed as male or female. In addition to gender, the visual aesthetics of artificial figures are especially emphasized – as beautiful (like Pygmalion's sculpture, Pandora, Olimpia, etc.), repulsive (Frankenstein's monster), or just unobtrusively human-like (Ning Schï's wooden man). These stories precisely describe the figures' intricately formed, lifelike details, and in course of this thesis we will understand why aesthetics are important. Responsive aesthetics lead to more initial acceptance by the viewer (*cf.* Hanson [108]) and seem to successfully obscure the artificial being. Because of their human shape, attractive, human-like figures can also be considered as a potential partner and sometimes produce sexual longing (as in the case of

Pygmalion, Pandora, and Olimpia). Combined with knowledge about a baleful motivation or a figure's unknown intentions, the eerie idea of the figure increases.

Cultural Context: Historical reports tell of the high level of craftsmanship and technical know-how necessary to create artificial humans. It seems worth mentioning by storytellers when the underlying technical processes of an artificial entity cannot be completely understood and wish to be like these characters and able to do things we cannot achieve by ourselves. In addition to the artificial entity itself as well as its creator, reason, materials, or the method of development are highlighted and depend on epoch, religion, and culture. In antiquity and the European Middle Ages, especially mystical or divine factors were accountable for the creation of artificial life. In Asia, in Arabia, and during the Renaissance in Europe, the art of engineering was especially highlighted. Today there is another image of robots in Asia (especially in Japan) than in the West. Thus, it is clearly significant in which culture an artificial figure exists.

2.1.7 Discussion

The presented aspects are relevant in forming both short and long term judgments of artificial entities and potentially influence the uncanny valley. Most of these judgments are rather negative, although there are a few positive examples such as Pygmalion's sculpture, the android Data from *Star Trek*, or the Tin Man from *Alice in Wonderland* that demonstrate a peaceful coexistence of natural and artificial beings. These promote the view that only equality, good intentions, and mutual respect can lead to higher emotional levels and relationships like friendship or love. A true example of this emotional bond is evident by the bomb defusing PackBot "Scooby Doo," which was mourned by US soldiers in Iraq after being destroyed by a mine explosion and could not be repaired [292]. Scooby Doo's case indicates that not only the appearance, but also the alleged common intention as well as its role in a group of humans may be critical for human acceptance.

The historical references often described artificial figures' life-like details. Why were these details as well as their skills and abilities so important, and why were they so precisely described? Artificial characters also need to look attractive and have significant aesthetic qualities. Did the narrators hope that their description would seem more interesting by mentioning these details? Assuming that these figures had really possessed all these abilities and properties, why, excluding the previously mentioned positive examples, were such figures unable to integrate permanently into human society? In other words, why did Olimpia attract adverse attention at the party? Consciously or subconsciously, the authors could create a sinister concept with their accurate description of artificial characters. These figures attracted attention due to their "life-like" details, but despite high visual aesthetics, artificial figures are unable to get the long term acceptance of humans. In the stories, they just bring evil upon the people, disappear, or get destroyed; A permanent and stable relationship between human and human-like figures is rarely mentioned.

The most plausible answer for the unstable relation between artificial entities and humans is that artificial entities were consciously or subconsciously exploited by their creators due to their uncanny effect, which increases the tension in their stories. Intention, aesthetics, and the cultural context are combined by the mind into a certain role model of a human entity. The historical review shows that not only observing a real figure but also the idea of an artificial human is sufficient to trigger eerie responses – not only towards human protagonists within the stories but also to the readers. This idea of an ambiguous creature can also be declared as uncanny like its real embodiment. Only a few people may have truly seen an artificial human in the past, but the conflict between man and machine seems comprehensible and plausible if the uncanny valley within our imagination is already taken into account.

Human-like artificial figure are potentially evaluated by the same criteria we use to evaluate humans. As this is the precondition for social interaction with anthropomorphic figures, we cannot ignore missing human attributes such as imperfections in facial expressions or errors in speech. This is evident in short-term processing of negative perceived sensory impressions, or in long-term notions and the resulting prejudices. Finally, the historical review shows that sensations as well as thoughts result in the same stereotypical image of artificial figures – imperfect and therefore negatively associated. To enable smooth social interactions with artificial characters humans must ensure that they are designed according to human expectations. Furthermore, we need to consider our understanding and image of artificial entities and ourselves. An entity's intention has to be clear, the appearance has to be appropriate, and we generally have to reject unjustified prejudices towards artificial entities caused by our historical or cultural backgrounds.

2.2 Frameworks of the Uncanny Valley

The existence of the uncanny valley was proposed when technological innovations were newly introduced and robotic devices were mainly prototyped to imitate human prosthetic body parts. However, Mori found that observers of realistic devices reported unpleasant feelings, which was not the case with simple designs looking less human-like. Mori's article [213] contains design guidelines for engineers of anthropomorphic robots but no empirically tested model or a clear, verifiable hypothesis. Nevertheless, the uncanny valley has become popular and topical in many domains. The uncanny valley has been applied not only to robots and prostheses but also to explain eerie effects of wax figures, dolls, and results of plastic surgeries to explain negative sensations. In the context of computer animations and human-computer interaction, the uncanny valley has often been associated with negative impressions of virtual humans. Reviews and critiques of computer animated movies or games with virtual humans, which "tend to suffer from imperfections" [145] on skin, lighting, eyes, or animation mention the uncanny valley in this context. Although Mori himself searched for answers to explain the phenomenon for robots, he states that more science is necessary to create a precise "map" of the valley in future work [213]. Thus, it is important to quantify the dimension of human-likeness and to measure affinity.

2.2.1 The Dimension of Human-Likeness

Investigations of the uncanny valley show some issues with a clear scientific evidence either confirming or rejecting the naïve hypothesis of Mori [145, 164]. This is caused by the vague terminology of the dimensions *human-likeness* as well as *affinity*. Researchers complain that the dimension of human-likeness can not be operationalized or mapped onto one single continuum but can vary in an infinite number of ways [39, 145, 192]. Considering this unspecific formulation one might assume that *any* manipulations cause negative affinity at a certain point of human-likeness, but as the following sections show, the uncanny valley only occurs in certain conditions. According to Kätsyri *et al.* [145], who reviewed multiple articles in the field of uncanny valley research, this is not the case. *Not every* kind of manipulation of the human appearance reduces human-likeness, and *not every* kind of manipulation elicits the uncanny valley.

Mori took multiple exemplary entities to explain different degrees of human-likeness and brought, for example, an industrial robot, a stuffed animal, a prosthetic hand, and a corpse into one single continuum (see Figure 1.3). According to previous work, neither the sum of human (arms, legs, face) nor non-human characteristics (material, surface) can objectively be operationalized as one single factor along one continuum. Consequently, the operationalized dimension of human-likeness has to be quantified by a subjective measure. Thus, this factor can be affected by the dependent variable (affinity) as well as unquantified factors. This means that different concepts can exist of *what a human actually is* and that these concepts interact with each other. Noteworthy conceptual examples related to the dimension of human-likeness are both realism as well as aesthetics, which are often influenced by attractiveness in humans and design properties in objects (*cf.* Section 2.3.5).

2.2.2 Measuring Affinity

Researchers found that there are not only problems with the operationalization of human-likeness. The Japanese term *shinwakan* also describes a rather vague concept without having an adequate English translation [16] and researchers potentially "may never come to a full consensus" [18] about its clear meaning. The latest translation of Mori's original article refer to *affinity*. However, there are also works that argue that *likability* is the most appropriate translation [18, 195]. Kätsyri *et al.* [145] argue that all translations either refers to the psychological construct of perceptual familiarity or to emotional valence. In contrast to the construct of perceptual familiarity, which only refers to known characteristics and qualities of humans and objects, emotional valance is a well-known and operationalizable psychological construct (*cf.* Kätsyri *et al.* [145] and Russell [259]), which covers any positive (likable, pleasant, attractiveness) or negative (eerie, unpleasant, aversion) emotional affinity. However, researchers found that concepts such as eeriness can be associated with other emotions such as fear or disgust [33, 122].

Subjective questionnaires such as the revised indices of the Godspeed questionnaire [17] developed by Ho and MacDorman [121] try to isolate and quantify these aspects using a broad conceptual coverage of aspects. The recent update of the questionnaire by Ho and MacDorman from 2017 [120] is to the current date (2018) the latest attempt to quantify human-likeness, affinity, and aesthetics in terms of uncanny valley research. According to other self-report questionnaires and as we learned in multiple studies of this work, some of the items in the questionnaire do not resemble typical or appropriate semantic differentials items. Additionally, the questionnaire does not include typical terms used in uncanny valley research such as familiarity or comfort and some indices were developed under the assumption that there is *a priori* a U-shaped graph between the dimensions of human-likeness and affinity. However, the questionnaire is a widely used measuring tool and considered “an important step forward in developing a common metric” [145].

The lacking operationalizability of the dimensions of the uncanny valley is a problem for researchers and there are a number of reasons that suggest that Mori’s hypothesis should not be taken literally. However, aspects of the uncanny valley aroused interest in different disciplines and Mori’s article has drawn significant attention in science and public discussions. A large number of studies were conducted to find evidence and explanations for the phenomenon. Some of these studies confirmed the predicted effects and potentially support Mori’s hypothesis under certain conditions (*cf.* Mathur and Reichling [195], McDonnell *et al.* [200], Poliakoff *et al.* [232], Piwek *et al.* [230], Yamada *et al.* [342], and Burleigh *et al.* [33]), but there are also studies with contradicting results – also only under certain conditions (*cf.* Seyama and Nagayama [287], Hanson [108], and Bartneck *et al.* [16]).

In the following overview of related work, we will demonstrate that the occurrence of the uncanny valley depends on the tested model and on which conditions it has been tested. Thus, the question *when* the phenomenon occurs is

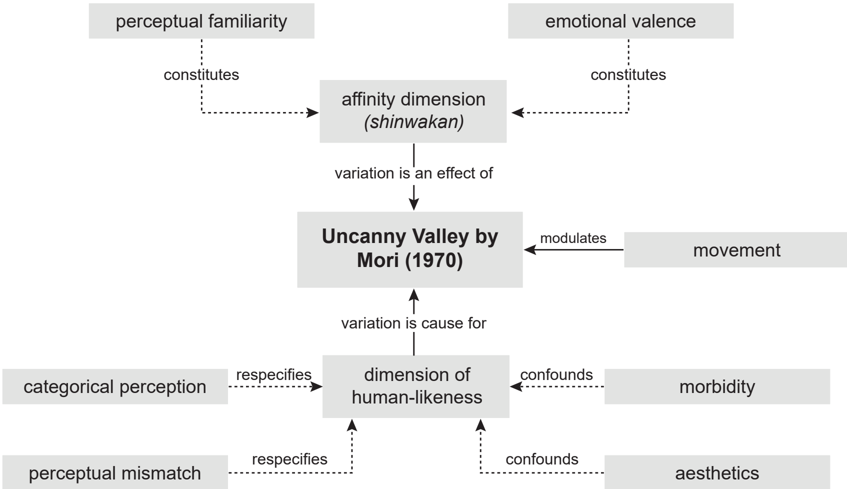

Figure 2.4: A framework map overview of empirically tested models showing the relations between the present hypotheses and the original hypothesis derived by Mori [213]. Dashed lines refer to cognitive constructs developed after Mori's publication in 1970. Illustration adapted from Kätsyri *et al.* [145].

directly linked to the question of *why* it occurs. We will show that the underlying cognitive mechanisms lead to an uncanny experience but none of them provides a *final* explanation of *one* phenomenon, which is *the* uncanny valley. In the following, we will introduce the current established cognitive models as well as proposed reasons that explain uncanny sensations in human-like entities. Kätsyri *et al.* [145] summarized these frameworks in a concept map. Figure 2.4 depicts an adapted illustration of this map.

2.2.3 Categorical Perception

To understand the mechanisms of human cognition, researchers often use perceptual conflicts. Perceptual conflicts can help to understand when and how people recognize, identify, or classify entities into categories that cannot be clearly assigned. This approach is the foundation of a systematic exploration of the uncanny valley based on the assumption that eerie characters evoke perceptual conflicts in the framework of categorical perception.

Originating in Psycholinguistics, *categorical perception* is the classification of stimuli into distinct categories when there is a gradual change in a variable along a continuum [99, 110, 244]. This means that similar physical stimuli either fall into one or another category and the transition is not being noticed. One example of categorical perception is that humans perceive and memorize words and their meanings but not the individual vocal sounds they make [171]. As a consequence of categorical perception, human subjects must already know or learn new categories. Then, humans discriminate between two stimuli most accurately when the pairs straddle the so called *category boundary*. Research has found that categorical perception is not only limited to a single low-level continuum such as auditory cues but is also applicable to other modalities such as haptics [90] and to high-level perceptual continua such as colors, faces, and emotions [20, 89, 250]. Generally, categorical perception allows us to distinguish between entities independently from how similar they are.

Categorical perception is a well-established construct in cognitive science. To demonstrate that categorical perception occurs, *identification* and *discrimination* tasks are used. In identification tasks, participants classify stimuli which exist on the continuum between two categories. Increased response times are used as an indication of categorization difficulty. In discrimination tasks, participants must decide if two stimuli (A and B) are identical or not. An extension of the AB-task is the ABX-task, where A and B must be distinct but one matches with a third stimuli (X). One approach in testing categorical perception is that the stimulus pairs of a discrimination task are derived from a foregoing identification task.

Categorical perception has been applied to human-like characters to investigate if the hypothesis of the uncanny valley is caused by difficulties in categorizing. Using identification and discrimination tasks studies empirically showed that morphed stimuli ranging from virtual to real humans are perceived categorically [40, 42, 179]. Yamada *et al.* [342] and Burleigh *et al.* [33] found that categorization difficulties are potentially associated with negative responses. However, there are serious concerns about these findings. According to Kätsyri *et al.* [145], it

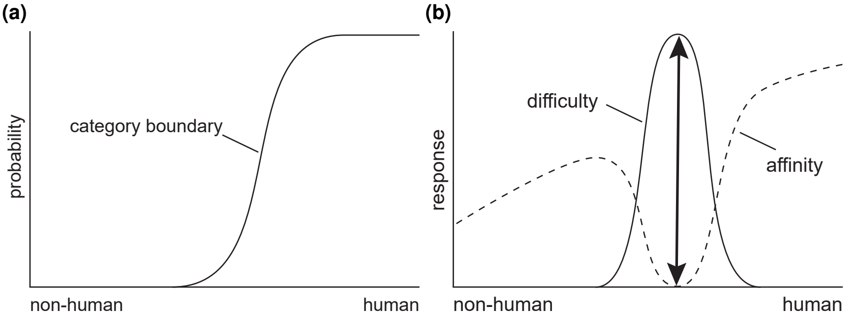

Figure 2.5: Difficulties in categorical perception. It is hypothesized [33, 40, 41, 43, 342] that difficulties in categorical perception (*e.g.* indicated by increased response times or lower ratings) along the category boundary (a) cause negative affinity between human and non-human entities (b). This hypothesis could currently not be confirmed [42, 145, 187].

is possible that artifacts in the morphed images comprised the human-likeness manipulation in the study by Yamada *et al.* [342] and no category boundary was found by Burleigh *et al.* [33], which threatens the validity of their study. Thus, none of the previous work investigating if categorical perception can be applied to the uncanny valley could show that characters located at the category boundary evoke more *negative* affinity than characters at the left or right side of the category boundary [42, 179]. Furthermore, previous work could not show that stimuli pairs straddling the category boundary are more difficult to distinct than other character pairs located on the same side of the boundary [42, 179]. Cheetham *et al.* [42] showed that difficulties in perceptual discrimination for adjacent stimuli pairs on the dimension of human-likeness are not associated with negative affinity. Thus, their findings indicate that the hypothesized relationship between categorization difficulty and affinity is “very likely wrong” [42]. The hypothesized outcome of the uncanny valley in categorical perception is illustrated in Figure 2.5.

2.2.4 Perceptual Mismatch

Initially not related to investigations of the uncanny valley phenomenon is the work by Feldman *et al.* [78]. Their work described a Bayesian model of optimal statistical inference that showed that stimuli which are close to a category boundary in categorical perception and judged by observers to be more dissimilar than stimuli away from that boundary. Investigations in categorical perception showed that humans perceive (virtual) faces categorically; however, as shown in the previous section, categorical perception seems not responsible for negative affinity predicted by the uncanny valley. Based on Feldman's Bayesian model, Moore [210] hypothesized that the uncanny valley is another manifestation of a perceptual distortion due to the misalignment of *multiple* category boundaries at the same time. Moore's model predicts that multiple features at different levels of human-likeness increase the probability that a misalignment between two (or more) categories occurs [210]. Moore's approach also showed that affinity between both categories can decrease due to the effect of *perceptual tension*. Perceptual tension occurs when "when the reliability of information derived from alternative cues to category membership is not balanced across different observation dimensions" [210]. In other words, perceptual tension occurs when an entity between a human and a non-human category has features that look more real than others. Figure 2.4 shows how the overlapping functions in the model cause negative affinity.

Moore's model is not fully compatible with the phenomenon of *perceptual narrowing*. Human infants are able to discriminate human faces and faces of other primates equally well. When they mature, they become better in discriminating faces of their own species and worse in discriminating faces of other primates [147, 225, 291]. This is supported by Lewkowicz and Ghazanfar [170] that showed that the uncanny valley effect emerges at 12 months. Using discrimination and rating tasks MacDorman and Chattopadhyay [39, 187] investigated Moore's model (see Figure 2.6) following the assumptions that perceptual narrowing influences the human "side" of categorical perception and that CG models are "inherently realism inconsistent" as the process of sculpting, texturing, and rendering of virtual characters automatically introduces inconsistencies in virtual realism. The authors revised Moore's equations and replaced the probability of occurrence in

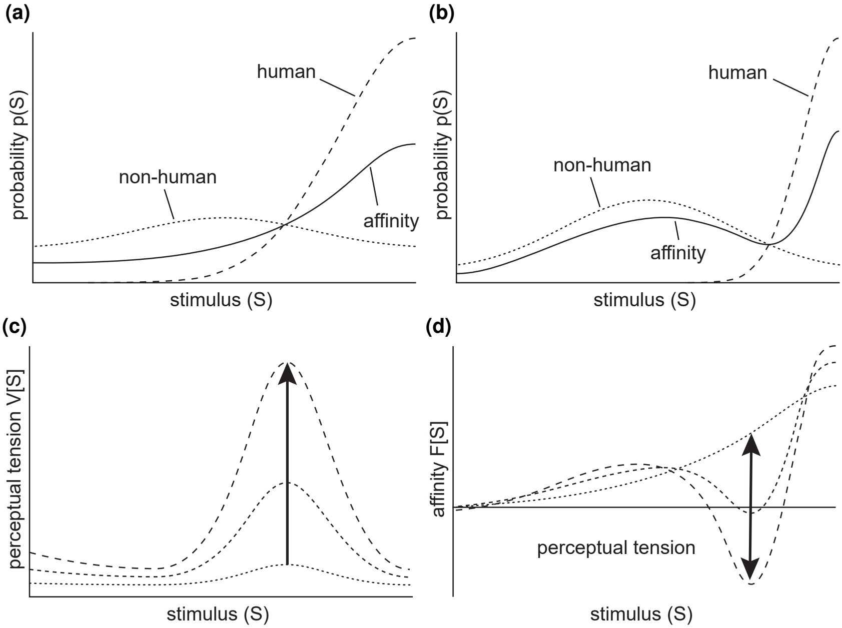

Figure 2.6: Bayesian explanation of the uncanny valley. The effect of perceptual tension introduced in the Bayesian model by Moore [210]: (a) and (b) show the probability of occurrence of stimuli with a broad *artificial* and a narrow *real* category. (a) shows a smooth transition between both categories leading to an almost linear relationship between the occurrence of value and the probability of categorizing a stimulus as real. (b shows that the relationship gets non-linear and monotonic when there is a smaller overlap between the categories (perceptual narrowing). The dip occurs around the category boundary at a smaller overlap between a real and an artificial category. (a) and (b) show the probability which cannot go negative and does not reflect the perceived eeriness. However, the perceptual tension increases at the category boundary as a function of uncertainty associated with *multiple* sensory cues as shown in (c). Perceptual tension leads to a dip in affinity and negative responses. The dip in (d) of the revised and empirically tested Bayesian model by Chattopadhyay and MacDorman [39] is related to perceptual tension and human-likeness which results for virtual humans in the characteristic curve of the uncanny valley (see Figure 1.3). The Bayesian model of Moore is given by the parsimonious combination function $F[S] = p(S) - k * V[S]$ with k as individual sensitivity. Illustration, description, and formula adapted from Moore [210], Chattopadhyay and MacDorman [39].

the Bayesian model with a subjective measure (perceived familiarity) to found support of an effect through perceptual narrowing using CG humans, animals, and objects.

MacDorman and Chattopadhyay [39] showed that negative affinity was not caused by categorical discrimination using stimuli along one continuum, however, by using multiple subtle transitions at the same time (facial regions or on image regions on artifacts). The authors found evidence that eeriness can be described as an inverse logistic function, based on the revised equations by Moore's Bayesian model [210] using perceptual tension as outcome of multiple conflicting sensory cues which do not look "entirely right" [145]. The more familiar an entity is perceived the more likely perceptual narrowing occurs. The perceptual tension increases, which caused the predicted (negative) affinity ratings by Mori [213].

In a further study, MacDorman and Chattopadhyay [187] showed that uncertainty and difficulties to identify to which category an entity belongs (human or non-human, living or inanimate, real or simulated) does not cause negative affinity. Instead, they showed that multiple unequal levels of realism cause uncertainty. Uncertainty about which features of an entity belongs to a category cause perceptual tension. Then the phenomenon of perceptual narrowing is responsible for the smaller human category in the Bayesian model of Moore [210] and explains why it is more difficult to create realistic looking renderings of human and animal features than parts of plants and objects. The authors showed that perceptual narrowing plays an important role in the cognitive development of facial familiarity and that this *perceptual mismatch* only occurs for humans and animals – but not for objects (see Figure 2.4).

Independently from these findings, further aspects of the perceptual mismatch hypothesis have been investigated [103, 187, 192, 209, 233, 287]. One aspect that should be highlighted is *atypicality* in human realism. In multiple studies, atypical features on faces particularly cause eerie effects [31, 192, 275, 287]. Examples of such features are realistic characters with unnaturally enlarged eyes [145], which were also intensively discussed in public reviews of movies and games with realistically designed characters in *anime* or *manga* style [323]. In context of the uncanny valley, this means that the visual system reacts more sensitively towards peculiarities in order to discriminate human faces. If some characteristics deviate while other features appear normal, this potentially leads to a perceptual conflict. The atypicality hypothesis refers to any distortions and deviant features from the human norm; however, currently only considers human-like characters [145]. In this thesis, we will later show how deviations and atypical features can also cause eerie sensations in familiar virtual animals (see Chapter 5).

Previous work found that atypical features are considered as most uncanny at very high levels of human realism [192, 300], which also supports the assumption that humans are less tolerant towards deviations of the average human appearance. Thus, research showed that both atypical features and inconsistencies in human realism are considered as artificial and reflect two aspects of the perceptual mismatch hypothesis [145, 192]. In their literature review, Kätsyri *et al.* [145] suggest that realism inconsistency as well as atypical features are potentially two different conditions of perceptual mismatch leading to the uncanny valley, as atypicalities were currently only investigated using human-like characters [145, 287, 300]. As indicated, our research in this thesis will show that atypicality also causes uncanny sensations using familiar virtual animal characters, which indicates that both aspects can also occur for virtual humans as well as animals and supports the assumption that both cognitive mechanisms belong to perceptual mismatch.

2.2.5 Inhibitory Devaluation

Current theories assume that Mori's hypothesis is directly related to human-likeness *per se*. Some researchers assume that the uncanny valley is a more general form of *devaluation* that occurs when *inhibition* is triggered. Inhibitory devaluation is an emotional evaluation phenomenon that generally occurs in any attentionally inhibited entity. Inhibition solves an internal conflict between competing entity representations. For example, it has been shown that inhibition effects a number of emotional judgments (*e.g.,* likability, trustworthiness, pleasure, etc.) [82, 87, 239] as well as the "incentive to seek and obtain otherwise appealing stimuli" [82]. The magnitude of inhibitory devaluation increases with the level of potential interference from competing stimulus-category or stimulus-response representations.

Negative consequences of inhibition were found in visual-recognition tasks which require classification–potentially linked to identification tasks in categorical perception without explicitly considering negative responses at the category boundary. Consequences of inhibition were also found for meaningless patterns [239], objects [105], and human faces [81]. Ferrey *et al.* [82] showed that negative affinity was not only caused between CG morphs of human-like entities but also using competing abstract stimuli (bistable images) of animals and objects, for example. Their study illustrates that perceptual conflicts are not only related to the dimension of human-likeness, and they predict that inhibitory devaluation is caused when multiple, competing stimuli are perceived.

In fact, inhibitory devaluation could be responsible for the uncanny valley and has also been considered by other researchers [146] as it explains negative affinity in categorization tasks. However, similarly to the study by Yamada *et al.* [342], the validity of the study could have been compromised by blending artifacts (see Kätsyri *et al.* [145]) through texture blendings on the 3D models in the second experiment by Ferry *et al.* [82]. Bistable images as used in the first experiment by Ferrey *et al.* [82] depict discrete entities of abstract objects and not a continuum of blurred transitions as in their second experiment. Even if the results were not compromised most aspects of this hypothesis could also be explained with the perceptual mismatch hypothesis and using the Bayesian model by Moore [210]. Moore pointed out that the perceptual tension in Bayesian explanation of conflicting stimuli is not limited to phenomena such as the uncanny valley. Perceptual tension would also cause negative responses such as disgust, for example, towards food that is off [210].

We summarize that difficulties in categorical perception do not explain the uncanny valley, but multiple conflicting cues in categorical perception trigger a perceptual mismatch, which causes the predicted negative affinity ratings. Thus, Moore's Bayesian model [210] solves the incompatibility between category uncertainty and perceptual mismatch. Studies have found good explanations for perceptual mismatch as underlying cognitive mechanism; however, there is currently no final explanation *why* the uncanny valley occurs.

2.3 Evolutionary Origins

2.3.1 Morbidity

The effect of perceptual narrowing in the study by Chattopadhyay and MacDorman [39] shows that there is a cognitive development of face recognition in early childhood; however, some researchers go further and assume that the uncanny valley has a genetic background and is a natural phenomenon in all primates. Steckenfinger and Ghazanfar [297] found that macaques showed less eye fixations while looking at realistic CG renderings of their species, however not for photos or unrealistic renderings. They conclude that there is an evolutionary origin behind the uncanny valley. Also Mori [213] assumed that the uncanny valley is caused by the need for self-preservation. Researchers discuss two mechanisms that ensure self-preservation: *pathogen* and *threat* avoidance.

The detection and avoidance of infected individuals are essential mechanisms against diseases and in reducing the risk of contamination [23, 260, 264]. Axelsson *et al.* [8] showed that facial cues such as pale lips and skin tones, hanging eyelids, or red eyes characterize a sick person. The authors also showed that the visual system in humans can detect even small indicators of sickness from observing human faces only a few hours after infection. It is assumed that there are more disease-specific features that developed over time to detect different diseases, and that this mechanism is potentially helpful in preparing the immune system. For example, Schaller *et al.* [265] showed visual cues of disease symptoms modulate aggression of immune responses. Pathogen avoidance is potentially also related to Rozin's theory of disgust [256], which posits that mammals evolve a disgust response to visual defects indicating that another subject is infected by a disease. Related to humans, Christian Keysers (cited by MacDorman *et al.* [192]) assumes that "the more human an organism looks, the stronger the aversion to its defects, because (1) defects indicate disease, (2) more human-looking organisms are more closely related to human beings genetically, and (3) the probability of contracting disease-causing bacteria, viruses, and other parasites increases with genetic similarity. Thus, leprosy looks disgusting to us, but leaf spot does not". The pathogen avoidance theory would explain the sensitivity to visual defects, which is much stronger towards members of one's own species than to defects in others [192].

Threat avoidance is directly related to the pathogen avoidance theory: Some diseases such as rabies cannot only be a disease vector but depict an immutable threat of the own existence. Yamada *et al.* [342], for example, assume that the brain categorizes an entity into a novel class and as a stranger which should be avoided. However, MacDorman and Chattopadhyay showed that *categorization-based stranger avoidance* is not responsible for the uncanny valley effect [186]. However, any avoidance behavior has a self-preservation effect and can summarized in a *morbidity hypothesis* as suggested by Katsyri *et al.* [145]. They authors also assume that morbid characters such as corpses and zombies mentioned in Mori's hypothesis "not suprisingly" elicit negative responses and *a priori* conclude that "it is quite trivial that such characters should evoke negative affinity" [145]. Of course, a fear of zombies is not plausible, but the aversion to them may exist due to an evolutionary origin, which remains unknown. Currently, there is no work that is able to link the aforementioned cognitive models with the morbidity hypothesis. Also Chattopadhyay and MacDorman [39] assume that threat avoidance is an evolutionary explanation for the uncanny valley; however, it remains unclear whether or how these feelings differ from emotions that trigger uncanny and strange sensations.

2.3.2 Mortality Salience

Mortality salience describes the awareness of the inevitability of death. MacDorman [183] assumes in an early work that human-like androids could subconsciously remind human observers of the fact that they too shall die. This reminder potentially triggers a so called *terror management defense*; an inherent refusal to lose cultural value, self-esteem, and individual traits [104]. The fear is caused by losing a buffer for existential anxieties and triggers terror. It is potentially caused while observing an android, which reflects death's inevitability and subconscious anxieties of reduction, replacement, and annihilation. For example, uncontrolled body movements of an artificial entity could elicit a fear of losing control of the own body. This reflection may be translated into an eerie feeling. This framework is potentially related with the morbidity hypothesis, however, it remains unclear which mechanism or which aspect might be responsible for uncanny sensations at high levels of human-realism.

2.3.3 Empathy

Another explanation for the reflection of subconscious fears could be empathy [208]. It has been shown that empathy is caused by mirror neurons in the human brain [91, 151, 248]. Furthermore, it has been found that animated human-like robots cause suppression effects mainly in the parietal cortex, which is known to contain clusters of mirror neurons [262]. Using such "shared circuits" [192] primates are able to understand intentions of others. Using functional magnetic resonance imaging (fMRI) Krach *et al.* [155] found a relation of human-likeness and cortical activation in the medial frontal cortex as well as the right tempo parietal junction, which shows that the human brain tries to reflect the intention of a robot. To understand intentions of others is a key feature in empathy. Furthermore, it has been shown that facial expressions and human appearance increase the ability to simulate of feeling the emotion of someone else [51, 236]. However, the explanation of lacking empathy does not explain why particularly almost human-like depictions trigger unpleasant feeling. Feeling empathy already begins at very low levels of human-like characters [292] and feelings in the uncanny valley seem not to interfere with empathy in perspective taking [190].

2.3.4 The Sorites Paradox

Similar to other researchers, Ramey [238] assumes that the uncanny valley is caused by the overlap of human and non-human categories. However, he also suggests that implausible features in human-like entities challenge the intuition of one's individual identity and humanity. This means that the confrontation with a realistic human-like character prevents a human observer from believing that the perceived entity is still real. Ramey [238] assumes that this disbelief occurs to indicate potential flaws and to stop perception from being tricked. He explains the emergence of this unbelief using the paradox of the heap (sorites paradox). For example, removing a single grain in a heap of sand does not matter. Following this consideration, removing the next single grain again and again would not significantly change the appearance of the heap. However, at a certain point, the heap changes in size and appearance so much that it is different from the original heap or becomes no heap anymore – the paradox occurs when only "large" and no "small" differences are noticeable and no mental differentiation between "small" and "large" exists.

Ramey [238] adapted the heap paradox as explanation for the uncanny valley and pointed out that only the sum of human qualities shape a full human being. If there are enough missing human properties, the "heap" no longer appears as a human, which causes the mentioned disbelief. Similar paradoxes were subject in color and sound perception and are related to ambiguity decisions as well as the category boundary in categorical perception. As we have learned, categorical perception has extensively been investigated in uncanny valley research, however, the existence of a perceptual disbelief within this framework is currently unknown.

2.3.5 Evolutionary Aesthetics

Evolutionary psychology assumes that aesthetic preferences of humans have evolved to enhance survival and to ensure reproduction [69]. According to Charles Darwin [57] physical attractiveness and the sense for physical beauty is an evolutionary mechanism for optimal mate selection and survival conditions. However, different cultures and epochs have different ideals of beauty and attractiveness, which is partly compatible with Darwins explanation of attractiveness [135]. Findings in neuroscience indicate that sexual and aesthetic preferences are two different processes in the human brain. While sexual attraction is processed by the limbic system, consciously aesthetic judgments are mainly made in the cerebral cortex [131, 228, 251]. This indicates that reactions to a biologically attractive stimulus run usually unconsciously, while aesthetic judgments are comparative and consciously mental decisions.

Physical attractiveness is perceived with averageness, symmetry, and sexual dimorphism, which are universally considered as appealing [159, 227, 245, 309]. Investigations of categorical perception or in difficulties in categorization often use morphed images to produce an objective dimensions of human-likeness. For example, the results of the third experiment conducted by Yamada *et al.* [342] indicate that their results contradict the assumption that difficulties in categorization cause negative evaluation. Morphed images between two males showed highest likability ratings at the point of the most ambiguous categorization. It is assumed that average faces not only have ideal proportions, but the computational blending of images causes a smooth skin structure, which is considered as one of the most important factors in human attractiveness [345].

As there are aesthetic preferences for both human as well as non-human entities, all aspects along the dimension of human-likeness in Mori's uncanny valley are potentially affected by aesthetics. Mori's article indicates that purposefully designed entities can overcome eerie sensations [213]. Hanson [107] assumes that designers of robots only have to increase the quality of their aesthetic design in order to avoid the uncanny valley. However, aspects of attractiveness of the human category have been little studied so far. We will learn in this thesis (see Chapter 4) how physical attractiveness may help to overcome eerie effects on the human side of the uncanny valley [278].

Researchers found that there is a relationship between pathogenic prevalence and aesthetic sensations in mate selection [92, 180, 309]. Visual cues that indicate parasite and pathogen resistance are perceived more attractive, which depends on the health of the population. People in geographical areas with more pathogen incidence value physical attractiveness more than people in areas with little pathogen incidence [92]. This indicates that the theories of morbidity and aesthetics are potentially part of the same evolutionary reason for the uncanny valley: preservation of the own existence and the own species, and to distinguish suspect human or animals entities, especially in cases where their condition and appearance indicate a potential threat. This theory is supported by multiple studies in the field of uncanny valley research, which showed that human-like entities trigger avoidance behavior [195, 300, 301].

Theories in evolutionary aesthetics not only regard physical attractiveness but all kinds of aesthetic sensations such as for objects, animals, and environments. It has been found that humans respond positively to any aesthetic stimuli that have been conducive to survival, reproduction, and heredity of the own genes. For example, studies have shown that people in all cultures consider river landscapes as well as open forested landscapes as visually appealing [223]. This may be due to the fact that early human cultures living in steppe or savanna relied on flowing waters and trees that provided food and shelter [327]. This example shows that also the non-human side of the uncanny valley is affected by aspects which can potentially be explained by evolutionary aesthetics.

Related to aesthetics in the uncanny valley hypothesis, Hanson [107, 108] assumes that any level of realism can be affected by the uncanny valley and proposed another theory called “path of engagement” to overcome the uncanny valley. Hanson suggested that “good design” [107] can help to make every robot “loveable and part of the human family” [107]. Hanson assumes that the uncanny valley does not trigger fear but rather “surreal” feelings and that the people’s sensitivity towards aesthetics increases with the level of realism. The study by Hanson showed no negative affinity ratings of morphed cartoon and photo images, however, as pointed out by MacDorman *et al.* [192] and Kätsyri *et al.* [145] the morphing technique used in Hanson’s experiment could have compromised the eeriness ratings. Currently, there is no empirical evidence or further investigation for the path of engagement theory.

Cultural or regional differences in the perception of human-like androids were also discussed by previous research of the uncanny valley. One example, which is likely not related to an evolutionary but a cultural influence, is the affinity of people towards robots in Japan. MacDorman [191] showed that people in Japan and people in the US have very different attitudes or anxieties towards robots in one’s everyday life. Robots in the US are associated with threats to their jobs while robots in Japan are used for labor shortages in elder care or health institutes. This example does not necessarily show that evolutionary aesthetics influence the preferences of a culture, but that culture influences an individual’s preference and attitude towards robots and androids.

2.4 Effects of Movement and Interaction

Research investigating Mori’s uncanny valley [213] often uses static stimuli such as rendered CGI characters on stills as they are easier to produce and to manipulate than using physically existing robots. However, the original hypothesis of Mori [213] refers to the appearance of robots or prostheses that physically exist and *move*. In his article, Mori [213] suggested that movement of an entity amplifies the uncanny valley curve (see dashed line in Figure 1.3). This means that affinity decreases when an entity moves, but also decreases when it falls into the uncanny valley. In the meta review by Kätsyri *et al.* [145], only two studies were found that could be taken as explicit tests of the movement hypotheses [230, 308]. However, both studies showed that movement causes generally higher affinity and showed no amplification of the uncanny valley curve. Thus, both studies did not support Mori's assumption that movement amplifies the uncanny valley effect, but according to Kätsyri *et al.* [145], it remains doubtful whether both studies are valid tests of the hypothesis as neither was able to show a non-linear relationship of human-likeness and affinity for still characters as predicted by the uncanny valley [145].

That movements of humans, androids, and robots are interpreted differently by the human brain, was shown by Saygin *et al.* [263]. Using fMRI they found increased suppression effects in the bilateral anterior intraparietal sulcus; a key node for the action perception system in the human brain, while presenting videos of androids. The results potentially reflect a prediction error in the human brain that perceives a human-like entity which does not move as expected. Similar effects were found using electroencephalography (EEG) measures: two studies [217, 319] found a negative peak in the N400; an event-related potentials (ERP) component of the human brain which is associated with mismatched expectations. Furthermore, it remains unknown if Mori [213] considered movement as a function of human-likeness itself (less realistic motion for less realistic character, high realistic motion for high realistic characters) or as an independent factor that can be applied to any level of human-likeness. Both conditions have not been investigated yet.

However, researchers also studied interactions with robots. Findings by Zotowski *et al.* [348] showed that repeated social interaction with robots can improve affinity. A large set of stimuli with real-world photos of machines, robots, androids, and humans was used in a study by Mathur and Reichling [195]. The authors conducted a large-scale online survey and a social game to investigate the users' likability ratings and decisions related to the trustworthiness of the robots. The study revealed a detailed insight and a quantitative cartography of uncanny valley as well as the predicted curve by Mori [213]. The work by Mathur and

Reichling [195] not only shows an uncanny valley curve with photos of robots but also that robots as social interaction partners influence humans' reaction and behavior. These findings are supported by a study by Strait *et al.* [301], who found that there is a link between avoidance behavior and the eeriness towards human-like robots.

Early evidences of the uncanny valley in interactive video games for human characters were found by Schneider *et al.* [267] based on an idea and work by Duffy [67]. They investigated subjective ratings of participants towards images of game characters and found hints of a non-linear relationship of virtual characters in video games as proposed by Mori's uncanny valley [213]. Tinwell *et al.* [318] showed that animated facial expressions and human emotions may change the perception of virtual characters and also may cause the uncanny valley. The authors derived design guidelines suggest that designers should pay attention to animations of upper facial expression especially when trying to display fear and sadness. Tinwell *et al.* [315] also found that articulation and lip movement during speech influence the effect of the uncanny valley and that the detection of micro-expressions can be associated with psychopathic behavior [317]. Similarly, Mäkäräinen *et al.* [194] showed that exaggerated facial expressions of very realistic human depictions may also led to the uncanny valley. However, implications of a lacking affinity in an interactive virtual context are based on caveats in comparison with assessments in the real world [313]. Thus, research of interactions with virtual characters in previous experiments is still limited.

2.5 The Uncanny Valley in Virtual Reality

As indicated in the introduction, one key feature of upcoming VR technologies in HCI is rendering the user's body in the virtual world using avatars. Avatars in VR provide a natural and intuitive interface for the user to interact with the surrounding virtual world. The most important body parts for interaction using avatars are one's hands and fingers, which can be displayed in VR. Today's technologies allow motion tracking of hands and fingers without wearing additional motion controllers or markers. Thus, arms, hands, and fingers can be rendered in VR according to their real pose and location. However, prior work assumes that the uncanny valley applies towards the own body similarly to others. For example, a non-VR study by Poliakoff *et al.* [232] found that human-like prostheses of the limbs are rated as being eerie in contrast to mechanical or real human hands. Thus, some studies focus on the effects of realism on the perceived presence in VR using so-called hand avatars.

The aspect that avatars cause uncanny sensations at realistic levels is potentially related to the quality of immersion in VR. Brenton *et al.* [31] suggest that there is a relation between the uncanny valley in the feeling of presence – the feeling of ‘being’ and ‘acting’ a virtual environment. The effect on presence using different levels of realism was examined by Vinayagamoorthy *et al.* [324]. While being in a VR cave (surrounded by projections of the virtual environment), participants reported that they had the lowest degree of presence by using the realistic walking NPCs on a virtual street. In their study, the participants were present in the virtual world with their real body. Lugrin *et al.* [181] used hand avatars and measured a lower degree of body ownership using more realistic avatars in VR. They found a perceived elicitation of body ownership using non-human first-person avatars. Furthermore, they measured a lower degree of body ownership using more realistic avatars without regarding the degree of presence or immersion in VR. Authors of both works [181, 324] argue that their results are the outcome of the uncanny valley phenomenon [213]. An VR experiment by Stein and Ohler [298] showed that the eeriness of “digital minds” connects to threatened human distinctiveness. They found that empathy and social cognition in VR are proposed as particularly unique human traits. Furthermore, humans seem eerier when they demonstrate scripted instead of spontaneous empathy. Nevertheless, increased levels of avatar or hand realism seemingly leads to different findings. Using a first-person computer-game Christou and Michael [47] found that visual characteristics of the avatar influence the players’ behavior. This was confirmed by Argelaguet *et al.* [7] who found that hand realism has an influence on the sense of agency, which is stronger for less realistic virtual hands. However, the sense of ownership is increased for human-like virtual hands, which was shown by Lin and Jörg [174].

2.6 Summary

This chapter provides an overview of scientific research in the context of the uncanny valley. In an historical and cultural literature review, we explored that eerie sensations that could be associated with the uncanny valley occurred long before Mori coined the term. We found that narrators used the eeriness of artificial figures to increase the tension of their stories. In western culture, artificial figures were considered rather negatively, while in eastern cultures artificial figures have been regarded as friends or servants and their engineers as masters in crafts or even as artists. Historic reports of artificial characters often emphasize aesthetic qualities and their obscure or manipulated intentions. By exploring explanations (RQ1), we found that researchers of the uncanny valley assume that the confrontation with an almost realistic human challenges the intuition of one's individual identity and humanity [238].

According to Ramey *et al.* [238], the uncanny valley might be a mechanism designed to detect potential flaws in the human appearance in order to prevent perception from being tricked. Interestingly, the negative image of artificial figures is also apparent in the historical review. Many artificial figures were described as sinister or obscure, and their eerie appearance provide a hunch of their actual plans. The question remains, whether this hunch is innate or learned? Is there an effect in which reports reinforce this hunch or is this hunch the cause of these reports? In the remainder of this chapter, we looked at cognitive and evolutionary work by researchers who sought answers to these questions. Steckenfinger and Ghazanfar [297] found first support for the explanation that the uncanny valley has an evolutionary origin. MacDorman and other researchers conclude that the phenomenon evokes revulsion to avoid threats and pathogen infections [33, 39, 184, 185, 187, 192, 238, 318]. The gaze behavior of macaques indicate that the uncanny valley can be found in primates; however, their study was never conducted with human subjects. Thus, if the uncanny valley has an evolutionary origin, the behavior should also occur in humans and will be investigated in this thesis (Chapter 3).

Research on cognitive frameworks discusses which model predicts when the uncanny valley occurs. It has been shown that the uncanny valley as predicted by

Mori does not occur for every kind of human-likeness. First, researchers assume that difficulties in categorical perception cause negative effects [32, 40–43, 287]. In many cases, the framework was not sufficient to explain eerie and negative sensations [39, 186, 187]. However, researchers found evidence that *multiple* conflicting cues of inconsistent features in human realism cause discomfort. These conflicting cues arise while texturing, sculpting, and shading a virtual character [187] and evoke negative sensations for humans and animals, but not for objects [39, 186, 187]. Furthermore, researchers assume that atypicalities cause discomfort [145, 192, 287, 300]. However, negative sensation of atypicalities were only found for humans. Thus, it is unclear if atypicalities and inconsistent realism belong to the same construct, which is investigated in Chapter 5 of this thesis. Generally, an uncanny valley of animals was previously not examined by previous work (RQ3).

Interestingly, there is only little work investigating how the uncanny valley can be avoided. Using deduction regarding specific stimuli, researchers provide isolated guidelines for designers and developers of applications with virtual characters [192]. According to research in evolutionary aesthetics, physical attractiveness such as averageness, symmetry, and sexual dimorphism, are universally considered as appealing [159, 227, 245, 309] and could help to overcome eerie effects (RQ4) for almost realistic characters, which would otherwise fall into the uncanny valley. Furthermore, it is conceivable, that user defined avatars, could help to actively determine preferred as well as not preferred characteristics of virtual faces.

Some is already known about the effects of the uncanny valley on presence using avatars in VR. However, it is unknown if the uncanny valley in VR is triggered by the same mechanism as in the real world and how deviations from the one's own appearance fit into the dimension of human-likeness (RQ5). Mori [213] mentioned that the uncanny valley could be avoided using stylization, which could be confirmed by previous work [199, 200]. However, using stylization artists often present altered body structures such as four-fingered hands. In terms of virtual reality, it is unknown if reduced human-likeness without considering the user's body structure decrease presence in VR (RQ6). In the further course of this thesis, we also examine if avatar realism in virtual reality has an impact on the performance and the feeling of presence of the user in VR (RQ7).

3

Importance of Eyes

Eye-tracking is used to understand cognitive processing and visual perception. The following chapter investigates whether artificial figures in the uncanny valley affect human's gaze behavior. We use eye-tracking to understand how humans perceive facial features and change their scan paths or observation schemes towards human and human-like depictions. Furthermore, we investigated how human-like characters are assessed when they are real or artificial. To investigate morbidity fears we have also presented pictures of dead people. We found that eye gaze behavior significantly changes when a character is perceived as eerie. Our work emphasizes the importance of the eyes when inconsistencies in human faces should be avoided.

This chapter is based on the following publications:

V. Schwind and S. Jäger. "The Uncanny Valley and the Importance of Eye Contact." In: *Mensch und Computer 2015 - Tagungsband*. Vol. 15. Berlin, München, Boston: DeGruyter, 2015, pp. 153–162. DOI: 10.1515/9783110443929-017

V. Schwind and S. Jäger. "The Uncanny Valley and the Importance of Eye Contact." In: *i-com* 15.1 (2016), pp. 93–104. ISSN: 2196-6826. DOI: 10.1515/icom-2016-0001

3.1 Face Perception using Eye-Tracking

As indicated in Chapter 2, one of the most important questions in uncanny valley research is whether the phenomenon has an evolutionary origin. An evolutionary origin would mean that the uncanny valley is an innate process in humans and that people are implicitly affected. One approach to study evolutionary origins is to look into similar effects in human and other primates. Steckenfinger and Ghazanfar [297] found an effect of synthetic monkey faces on gaze behavior in macaques. The fixation times of their subjects were reduced using intermediate levels of realism of CG monkey faces. However, the study has not been conducted using humans and human faces before so there is no reason to assume *a priori* that the phenomenon also occurs in human subjects or that this effect is related to negative responses due to the uncanny valley. Thus, it is important to learn if similar effects occur in human gaze behavior and if this finding potentially supports or refutes current explanations of the uncanny valley (RQ1). Moreover, eye-tracking could give further valuable insights into human visual perception of virtual faces.

Eye-tracking has had a profound influence on our understanding of the mechanisms in recognition of faces and has been established as an important method of collecting empirical measurements [66, 142, 288]. Eye-tracking and area of interests (AOIs) on stimuli provide important information on the number and duration of fixations within areas of faces, for example. The disadvantage of this method is that variations can arise due to different patterns of AOI boundaries and individual differences possibly lead to inconsistent results. For this reason some studies reject such templates completely and focus on the differences between the stimuli [322]. It is noteworthy to mention that this method only works for figures with identical proportions of facial features. A wide range of stylized, drawn, uncanny, and real faces suggested by Mori's graph would not allow drawing direct comparison of gaze sequences.

As the literature review shows, eye realism seem to play an important role in the perception of faces. In particular, the "dead eyes" symptom [39, 143, 187, 272, 317] was mentioned to cause uncanny sensations. Therefore, special attention of this investigation was given to the eye regions of the stimuli and the gaze behavior of our participants. This chapter focuses on differences in eye movement behavior depending on the kind of human-like depictions observed while we mainly focus on fixations on the eyes region. Two further questions in this context are how eye contact is relevant for categorization of negatively rated characters and whether there are fundamental differences in observation schemes when photos of ordinary or CG humans are presented. Answers to these questions could provide a better understanding of human perception with artificial entities and the first hints on how to improve interactions with human-like entities.

The uncanny valley hypothesis refers both to the overall impression created by a figure as well as to prostheses. However, in the perception of human-like figures, increased attention is undoubtedly paid to the face and the eyes [112, 134, 137, 161, 249, 325]. Farah *et al.* [73–75] assume that face recognition is fundamentally different from object recognition. Even the simplest stroke patterns (*e.g.* emoticons) or templates can be interpreted as facial characteristics. Among all other facial features, the eye area draws the most attention and the highest number of eye fixations (40%), and thus shows the highest attention rate in the human face as shown by previous research [134, 137]. Previous work using artificial characters in uncanny valley research support this assumption [43, 83, 179, 192]. Mori associated the phenomenon with an aversion towards corpses and MacDorman [184] argues that eerie anthropomorphic figures trigger the subconscious fear of death. As indicated in Chapter 2, this is a highly discussed topic regarding the uncanny valley hypothesis [122, 208, 316].

3.2 Method

As previously mentioned, Steckenfinger and Ghazanfar showed that the uncanny valley also occurred in monkeys [297]. Further findings lead to the assumption that there is an evolution-related cognitive connection between the uncanny valley and the morbidity hypothesis (see Section 2.3.1). For this reason, we use eye-tracking using human subjects to investigate the effect of stimuli depicting virtual, realistic, cartoon, or abstract figures. Deceased persons standing upright with opened eyes were included to determine whether the phenomenon occurs due to morbidity cues. Such post-mortem photography was particularly common in the 19^{th} century and was supposed to produce a particularly spine-tingling effect on the viewer [201].

3.2.1 Study Design

An exploratory within-subject study was carried out to investigate the effects of virtual faces on the gaze behavior of our participants. We recorded eye movement and asked for subjective ratings of perceived realism, human-likeness, likeability, and attractiveness of each figure.

3.2.2 Stimuli

Sixty-eight images of persons and figures were created with 3D modeling software, captured from movies and games or chosen from the internet. Authors and owners of images gave consent for the use of their images in this study. The sample composition represents a cross-section of character styles that are related to the scale of human-likeness in Mori's graph and have been examined or mentioned in previous studies. Based on the type of face, the portraits were classified into these 8 categories: 13 real persons in photographs (ordinary human), 21 CG (CGI), 9 cartoon figures, 5 wax sculptures, 5 geminoids (androids), 5 humans with visible impacts of cosmetic surgery, 3 deceased (post-mortem), and 3 hyper-realistic cartoon figures with a real look. There were 11 pairs of images with the same or a very similar person and the same posture. Four images showed a pair: a person and his or her double. All depicted figures had a neutral facial expression and an upright posture. Computer-generated faces can be recognized due to junctions or contours. Cheetham *et al.* [43] noted that additionally to the mouth, eyes, and nose, other areas such as hair, hairline, and head contours might be of interest in an eye-tracking study. For this reason, we masked faces with no oval overlays and made no additional changes to the back- and foreground. None of the stimuli was manipulated. The only change applied to the figures was a uniform cut of the face sizes to fit a Full-HD screen resolution.

3.2.3 Procedure

Participants sat upright on a fixed chair in front of a 21.7″ LCD Monitor in a soundproof lab. Non-reflective whiteboards were used to prevent reflections of infrared light. The background brightness was 320 lux. Every procedure took about 20 minutes. At the beginning of the eye-tracking test, each participant was instructed about the procedure itself (calibration, first test, session). After 30 slides had been demonstrated, a participant took a break and received further instructions about the categorization task. The viewing distance was 60 cm. At the beginning of each experiment, a 12-point-matrix calibration was conducted with every participant. To avoid fixed gaze in the same position as on the previous slide, a black screen was shown between the stimuli.

All images were presented for 10 seconds. The subjects could move their eyes freely. After a random sequence in a group of 30 stimuli, a pause was followed by an instruction slide providing information about the upcoming reaction measurement. The subject had to press the buzzer if he or she believed that a figural representation was not a real human. Two participants asked what "real" meant. We repeated the task orally and added the note that some human representations were artificially produced. To prevent confusion between stimuli and the instruction slide, the test only continued after the participant made an initial input. Thirty-eight further stimuli of the categories ordinary human, wax sculptures, cosmetic surgery, and CGI were shown for 10 seconds. When a real figure appeared and the subject gave no response (did not press the buzzer), it meant the subject recognized a real figure correctly. When an artificial figure was shown and the subject gave no response, it meant that the artificial figure was perceived as a real human (and thus "passed" the test.) With this experimental scheme, it was not possible to measure response time when real human figures were displayed, since in this case participants simply had to wait until a trial finished. The benefit of this method was the possibility of getting precise measurements of the affective reaction with very low latencies when unrealistic figural representations were demonstrated. Multiple selections during a complex categorization task could have led to delays and long considerations.

The last task in this experiment was a categorization task with four final images. On every image, there was a real human and the same person with a digital double or a doppelganger android. The subject had to press the left or right keyboard button to identify the “fake.” This part of the experiment was not included in the regular AOI calculations or other fixation measurements of the previous stimuli. This was the final part of the eye-tracking experiment. The eye-tracking device was a Tobii X2-30 Compact Edition with 30 Hz. Recording and playback of the slides were carried out with OGAMA 4.3 [329] on an Intel i7 3635QM with 2.4 GHz and 32 GB RAM. After each recording, a backup of the eye-tracking data was copied to an external drive via batch script. After every eye-tracking session, a questionnaire was handed out. On a numeric rating scale from 1 to 10, participants were asked to state subjective value for realism, human-likeness, likeability, and attractiveness of each figure. Then the participants were asked to indicate whether they had already known the figure before. To specify additional characteristics of the presented figures, participants could freely choose between 15 positive and 15 negative attributes randomly assorted in a multiple-choice matrix.

3.2.4 Participants

Participants (N = 53, 26 male, 27 female) with ages ranging from 18 to 63 years (M = 31.7) were volunteers among students and staff of the University or visitors on the campus. All stated that they had no mental or physical illness. Twenty-one participants were wearing glasses during the examination. Forty five participants were German, 2 Chinese, 1 Indian, 1 Italian, 1 Mexican, 1 Pakistani, 1 Turkish, and 1 from the USA. Twenty one participants claimed to have no experience with computer and video games. Sixteen played video games once per week, 12 several times a week, four daily. Three participants never watched featured films, 15 watched one per week, 28 several times a week, and 7 watched them daily. There were 13 actives in art or humanities, 11 engineers, six social scientists, 15 involved in nature sciences, and eight non-academicians. Fourteen participants stated that they had heard the term “uncanny valley,” before. To get a profile of the individual view toward artificial figures, we asked the participants whether it makes sense for them to build or to simulate human-like figures. Sixteen gave their consent. Before the eye-tracking test, all participants declared their agreement to data protection and privacy policy. After the test, they were asked

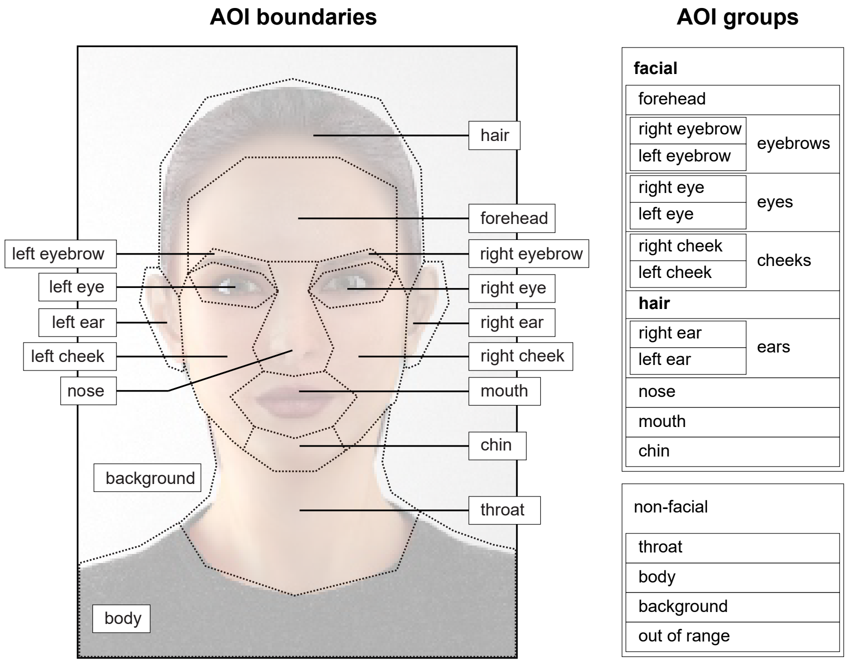

Figure 3.1: AOI boundary template for the facial regions. Predefined 16 AOI boundaries on one of the presented faces. Divided and symmetrical areas were summarized into AOI groups of respective pairs (eyes, ears, cheeks, ears).

to fill out a questionnaire. One far-sighted participant wearing glasses could not perform the test because of too strong diopter values (+12, r) that could not be calibrated by the device.

3.2.5 Data Analysis

The recorded raw data were aggregated to eye fixations. A fixation was calculated by the maximum distance of 20 pixels (.45°) and the minimal number of at least three samples (100 ms) of successive gaze positions. Lost data through eye blinking or fixation outside the screen were discarded. The first fixation of a stimuli recording was not deleted. AOIs, events, and the fixation table were exported separately for further analysis. The analysis and calculations of AOI hits were performed using SPSS V.21 and Excel 2013. To delineate AOIs, we used a template for polygonal boundaries shown in Figure 3.1. The template we used was derived from a preliminary study on the accuracy and precision of eye-tracking devices [275]. To understand the significance of facial features in the processing of eerie faces, we decided to use a common face mask and thus apply similar boundaries of polygonal regions of interest on each face.

It is important to note that AOI sizes had to be different due to the presented stimuli. However, we assume that we can make reliable statements about the proportion of attention within the specified AOI boundaries and with a relative calculation of facial fixations and dwell time. To be able to compare several samples with absolute fixation times, we provided identical AOI for 12 stimuli pairs – usually that of the same person or character (cartoon vs. hyper-real, CG vs. real, etc.). Fixations outside the predefined regions, by Figure 3.1, were treated separately (*e.g.*, neck, background, out of the display). The AOI boundaries were invisible for the participants.

3.3 Results

3.3.1 Questionnaire

Subjective assessments design the sorting of the categories in graphs and descriptions of the human-like axis in further diagrams of the study. The resulting ratings follow the order of human-likeness: (1) Hyper-realistic cartoon ($M = 1.667$, $SD = 1.062$), (2) Cartoon ($M = 2.560$, $SD = 1.984$), (3) Cosmetic surgery ($M = 5.571$, $SD = 2.866$), (4) Robots ($M = 6.548$, $SD = 2.847$), (5) Post Mortem ($M = 6.976$, $SD = 2.540$), (6) CGI ($M = 7.857$, $SD = 1.905$), (7) Wax sculpture ($M = 8.786$, $SD = 1.337$), and (8) Ordinary human ($M = 9.119$, $SD = 1.419$). The sorting of human-likeness organized the subsequent evaluation of the categories without a linear morph continuo. The average subjective ratings of human-likeness and realism per trial show a strong positive correlation between human-likeness and realism ($\rho = .966$, $p < .001$, $N = 64$, $CI\text{-} = .011$, $CI\text{+} = .008$). Both realism and human-likeness result in the same categorization sequence (see Figure 3.2).

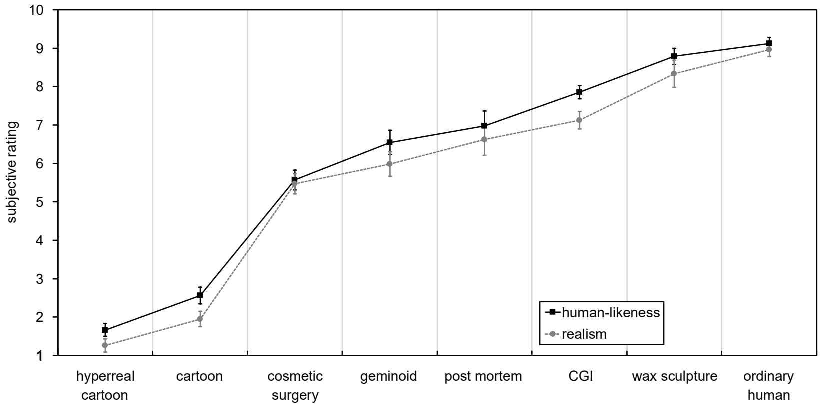

Figure 3.2: Ratings of human-likeness and realism from 53 participants. Ratings are sorted by increasing subjective assessments of the participants. Error bars show standard error (SE).

Attractiveness and likeability also show a strong positive correlation ($\rho = .921$, $p < .001$, $N = 55$, $CI\text{-} = .024$, $CI\text{+} = .018$). Ratings show differences between familiar and unfamiliar faces, especially in case of cartoon and hyper-realistic cartoon figures. Cartoon figures were rated more positively when they were familiar to the observers. An independent t-test shows a significant difference in the ratings of known ($M = 8.200$, $SD = 2.203$) and unknown ($M = 4.562$, $SD = 2.257$) cartoon figures, $t(80) = 7.233$, $p < .001$.

The examination of a multiple choice matrix with 30 different characteristics as a supplement of the subjective scale ratings provides additional insights: at first glance, the participants were certain about the category they could attribute to the observed figural representations. Deceased, for example, were categorized as artificial by only 6.98% of participants and as ordinary humans by 5.81%. In contrast, cartoon characters (46.51%), hyper-real cartoons (53.49%), and humans with visible impacts of cosmetic surgery (51.16%) were all often rated as artificial. Participants rated certain images as repulsive: hyper-real faces (30.23%), images in the cosmetic surgery group (34.11%), and the deceased (60.47%). Interestingly, 44.19% of the participants assessed neutral faces of the dead as "aggressive."

The same holds true for hyper-real cartoons (34.88%) and cosmetic surgeries (27.91%). This possibly indicates that the deceased was immediately perceived as a threat despite their neutral facial expression. This phenomenon is illuminated in detail in the analysis of the eye-tracking data.

3.3.2 Response Times and Fixation Sequences

To answer our research questions and to facilitate the analysis of the eye-tracking data, it was necessary to clarify whether significant changes in gaze behavior occurred during the reaction test. In this part of the eye-tracking test, the subject was asked to press an input button when the presented figure was not a real human. We found no differences between the relative fixation times on a face. An equivalent test confirmed this result: we performed a paired t-test, which showed no significant differences within the relative distribution between the dwell time in rated and unrated trials ($p > .05$), except for the forehead region ($p = .007$). The same analysis of fixations before and after the identification of a virtual figure was conducted within identification tests. This analysis provided similar results. During the reaction test, no significant differences between the ratios of facial attention were recorded.

To examine whether the task had an impact on the overall number and duration of fixations, we considered the average number of fixations and duration before and during a reaction measurement. An independent t-test showed that the average duration of a fixation before ($M = 354.87$, $SD = 370.72$) and during a reaction test ($M = 350.49$, $SD = 358.53$) differed with a mean of 4.389 ms and was minimally elevated (higher attention) with no significant difference between the two, $t(77032) = 1.716$, $p = .086$. We assume that the reaction test had apparently no significant effect on the number and duration of fixations. Since the subjects were not informed that a check for fakes of the figures would be part of the test, we can exclude the possibility that this was of any influence on their gaze behavior. However, we can assume that participants knew that it was part of the experiment. Such conscious analysis of artificial figures may take place in reality as well. We therefore assume that subjects searched for the truth of whether a character was genuine or not, and thus displayed no concrete changes in gaze behavior. However, the relative dwell time on facial features varies

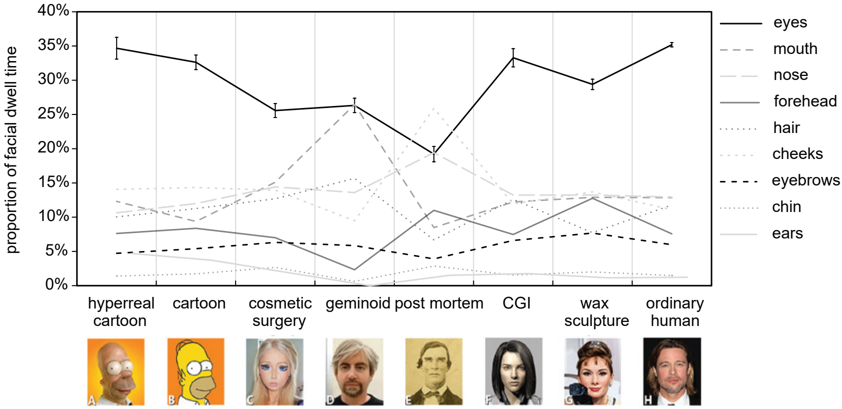

Image Credits:
(A) "Untooned" Homer Simpson, (c) 2008 Jax Pixeloo / pixeloo.blogspot.com
(B) Toon figure Homer Simpson, (c) 2000 Matt Groening / Fox Broadcast
(C) "Real Life"-Barbie., (c) 2012 Valeria Lukyanova / facebook.com
(D) Geminoid DK, Screenshot from the video Geminoid DK Mechanical Test, (c) 2000 Aalborg University / c.aau.dk
(E) CDV Man Post Mortem, (c) 2000 anyjazz65 / flickr.com
(F) CG film figure "Aki Ross", Final Fantasy - The Spirits Within, (c) Square Pictures / Columbia Pictures
(G) Audrey Hepburn's wax figure at Madame Tussaud's Sidney (c) 2012 Claudia Betia / flickr.com
(H) Brad Bitt, (c) 2012 s_bukley / shutterstock.com

Figure 3.3: Relative dwell time on facial regions. The curves show the ratios of the average dwell time on various facial regions. The error bars on the eyes curve show the standard error (SE).

between the categories. Figure 3.3 illustrates a significant decline of fixations and dwell time on the eye regions of figures that were rated as less attractive and had an average human-likeness. Depending on the category, attention was paid to the surrounding features instead. In figures after cosmetic surgery, fixations shifted to other areas of the face, whereas in geminoids the attention tends to be given to the mouth region. A difference of 1.1% between the dwell time on the eyes of ordinary people ($M = 35.19\%$, $SD = 21.79\%$) and CGI characters ($M = 33.29\%$, $SD = 20.23\%$) is relatively low and exceeds no general significance level, $t(1800) = 1.776$, $p = .076$).

The largest differences between facial fixation time (15.98%) between the stimuli were recorded between the eye regions of ordinary people ($M = 35.19\%$, $SD = 21.79\%$) and post-mortem photographs ($M = 19.21\%$, $SD = 19.64\%$), $t(846) = 8.417$, $p < .001$. The subjects looked at the underlying cheek regions instead. A similar observation regarding menacing facial expressions was made by Moss *et al.* [203]. Their study pointed particularly to differences between

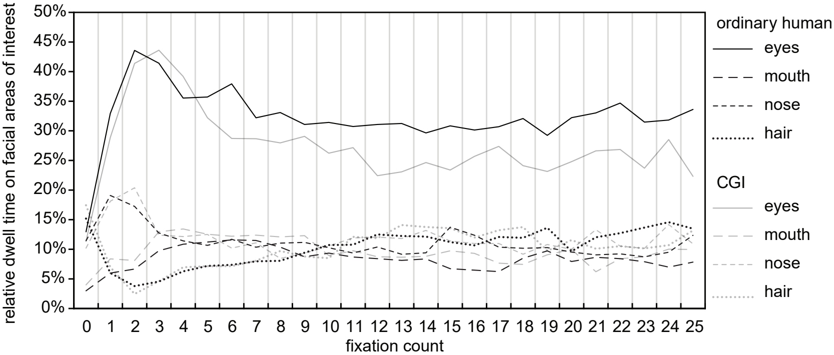

Figure 3.4: AOI hits of the first 25 fixations. The diagram shows the distribution of attention in four selected categories: eyes, mouth, nose, and hair for realistic and virtual characters (CGI). In the 2nd fixation (3) after the emergence of the image (1), almost 40% of all fixations landed on the eye region.

genders. A one-way ANOVA of all fixations within our study shows no significant differences between the occurrences of male and female fixations except for mouth ($M_f = .101$, $M_m = .084$, $F(1,51) = 5871$, $p = .190$) and background regions ($M_f = .130$, $M_m = .164$, $F(1,51) = 8585$, $p = .005$). These occurrences indicate that women fixated more frequently on mouth regions and men on background structures. As mentioned above, images of the dead were often rated in the questionnaire as aggressive and repulsive. No significant difference (.19%) was observed between the average facial fixation times of ordinary humans and CG figures ($M = 33.38\%$, $SD = 20.32\%$), $t(256,630) = 1.757$, $p = .076$. The analysis indicates that the most and longest fixations were recorded in the case of photographs of ordinary humans. Our evaluation demonstrates that eyes receive the most attention in all categories of human-like figures.

Figure 3.4 illustrates the relative dwell time on the four most important facial AOI during the first 25 fixations. At the very beginning of the trial, the eye regions attract attention like a magnet. The fixation sequences in all categories support this finding. Up to the 2nd and 3rd fixation, almost 45% of all fixations hit the eye regions. Also, nose and mouth were targeted increasingly at the expense of other features. For example, hair, chin, and cheeks played a subordinate role in the 1st fixations. After the initial visual contact (the 4th or 5th fixation), the importance of the eyes decrease. At this point, we can recognize a difference in the gaze behavior depending on the category. The relative attention paid to the eye regions of CG characters is slightly lower than that of photos showing ordinary people. This explains a small but measurable difference in the dwell times between the two categories in Figure 3.4. The difference is particularly obvious between the eye regions and can be explained by the fact that other facial regions of CG characters need to be examined more than for real people. The ambiguity provoked by the CGI faces results in fewer fixations on the eye region, and thus additional facial features must be considered to make a decision. The proportion of attention paid to these additional characteristics differs depending on stimuli (see Section 3.3.4). In the case of artificial faces, the distribution of fixations moves to other facial features at the expense of the relative fixation count on the eye region. However, in trials with both CGI and real figures, the most attention was paid to the eyes.

The graphs in Figure 3.5 show the occurrence of gazes resting on the eye region during the whole trial sequence within a time interval of 100 ms. The curves illustrate the relative occurrence of gazes focusing on the eye region of CG characters, cosmetic surgical patients, wax sculptures, and ordinary humans. Noticeably, a considerable increase of attention paid to this region is registered within the first 400 ms. The increase reaches its peak after 400–600 ms. After that, the attention decreases and settles at a stable level after about 1200–1400 ms (the time of the 4th and 5th fixation) until the end of a trial. Fixations on the eye regions reach a stable level in a time range in which most decisions have been made. With minor deviations depending on the degree of realism, this judgment applies to all categories, including photographs depicting ordinary humans.

When a CGI figure was demonstrated, 39 participants responded within 800–1000 ms. Additionally, 38 participants responded within 1000–1200 ms. When wax sculptures (16) and cosmetic surgery patients (12) were shown, most responses also followed within 1000–1200 ms. A period of 400–600 ms elapsed between the highest amount of fixations on the eye regions of CG characters and the average response time of the participants. As mentioned above, some real faces of ordinary humans and cosmetic surgery patients were rated as unreal. Overall,

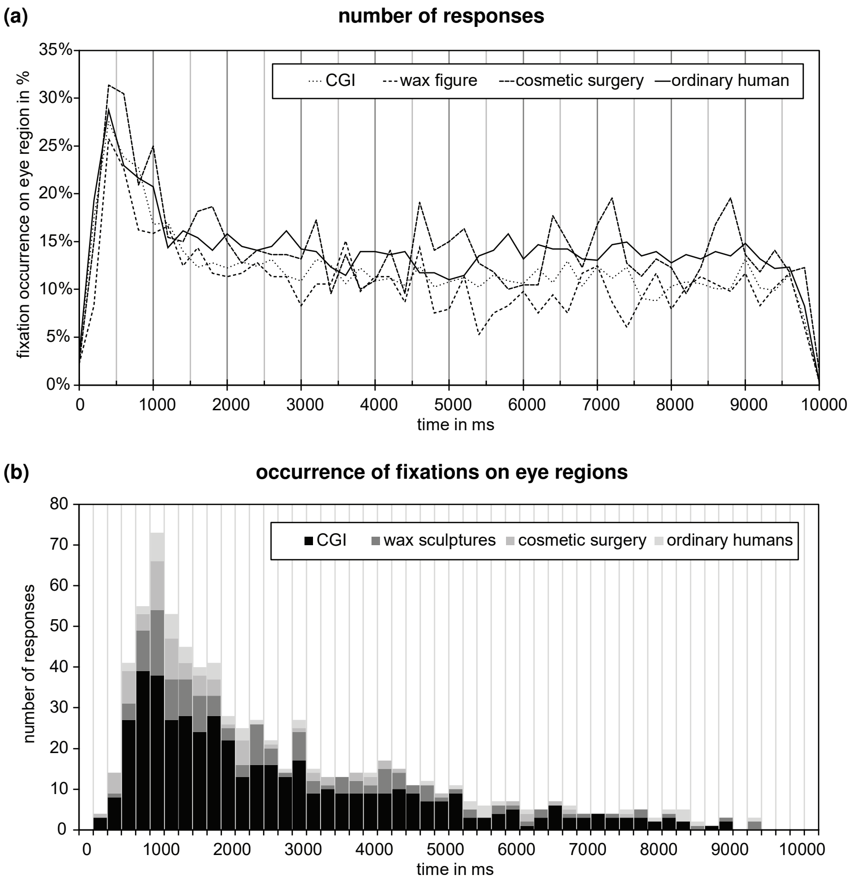

Figure 3.5: Results of the temporal gaze sequence. The upper diagram (b) shows the relative occurrence of fixations on the eye regions. The lower diagram (a) shows the relative amount of responses after stimuli onset.

87% of responses in the case of images of ordinary humans and only 36% of responses in the case of cosmetic surgery patients got a correct rating. Wax figures ended up with 50%, whereas CG characters had 66%. This explains the results of the questionnaire, according to which CG characters were considered unrealistic and less human-like than the wax figures, and were recognized more often. The high error rate in the case of images of faces notably altered by cosmetic surgery is evident due to their substantial deviation from a regular human face.

A certain amount of time passes between the first eye contact and the response. This suggests that eyes might be a significant feature for the assessment, but it does not clarify whether they are a specific trigger for the reaction itself. It is conceivable that during the last fixation, other regions attract attention before a reaction occurs. To find strong differences between the respective regions, we conducted a one-way ANOVA of the relative count of fixations. No significant differences could be found between the 4 fixations (eye regions: $F(3,208) = 1.477$, $p = .335$, all other: $p > .34$). Only trials with four fixations before a reaction were included in the analysis. The proportion of attention in the last four fixations before a response was given shows no significant changes.

The immense importance of the eyes in decision making is illustrated by the following results: in 78.45% of all trials (1580 of 2014 samples), participants fixated on the eye region at least once before they made a decision. If we consider adjacent areas to the eye region (eyebrows, cheeks, nose), these were 88.98% (1792 of 2015 samples). The mouth region was fixated in 59.15% and the nose in 69.98% of all cases at least once before the decision was made (hair: 47.61%, forehead: 46.05%, chin: 18.18%, ears: 11.12%). The importance of eye region becomes particularly evident when we count facial locations during the last fixation before a response follows. In 37.43% of all cases (280 of 748), the last fixation was in the eye region. Other AOIs were fixated significantly less (nose: 14.44%, mouth: 11.23%, cheeks: 12.3%, hair: 9.89%, eyebrows: 6.95%, forehead: 5.21%, chin: 1.74%, and ears: .8%). Thus we can conclude that eye-tracking data prove eye regions to be the most important criterion in the decision making process whether an artificial character is real or not. Eyes are also the most important facial feature regarding the proportion of attention when artificial figures are observed. Other facial features may also play a considerable role in unmasking and subconscious recognizing of artificial figures, but in most cases, no decision is made without fixating on the eyes.

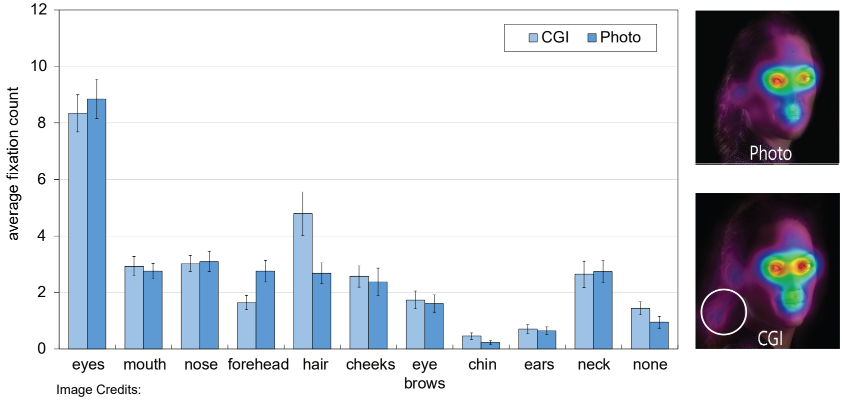

Image Credits:
(CG) Portrait of Marion Cottilard, (c) 2013 Nicolas Gaul
(PHOTO) Marion Cottilard, (c) 2009 by Anne Ryan / anneryanphoto.com

Figure 3.6: Direct comparison of the gaze behavior. Attention maps for a photo and a rendering of the same. The diagram compares the relative dwell time spent on individual AOIs in direct comparison. Error bars show standard error of the mean.

3.3.3 Correlations

Pairs of average response times and the mean of the respective subjective evaluation per trial revealed a strong positive correlation between the duration of a response and the specified realism of a figure ($r = .740$, $p < .001$, $N = 55$, $CI- = .070$, $CI+ = .054$). In contrast, reaction time and human-likeness correlate moderately ($r = .673$, $p < .001$, $N = 55$, $CI- = .083$, $CI+ = .066$). The later the response was given, the more realistic or human-like an image was rated. Attractiveness correlates weakly with the response time ($r = .383$, $p = .004$, $N = 55$, $CI- = .123$, $CI+ = .108$). The moderate correlation between likeability and the fixation time on the eye region ($r = .435$, $p = .001$, $N = 55$, $CI- = .118$, $CI+ = .102$) is the strongest correlation between the fixation times on facial areas and subjective ratings. Other significant correlations ($p < .05, r > .2$) between response times and subjective ratings or fixation time on facial regions could not be found.

3.3.4 Direct Comparison

Calculating direct differences in gaze behavior, we could identify some interesting, potentially "treacherous" areas of unreal characters. Compared to previously obtained findings, these were revealed less often. Our analysis shows differences in facial preferences of the observers before, during, and after the verification process. Local shifts of attention and absolute dwell time can be determined by a specific templates (*e.g.* CG or photo). While drawing a direct comparison between the attention areas in the case of an ordinary human image and a CG image, we noticed a significant shift of attention within facial areas when unrealistic features were observed. For example, Figure 3.6 illustrates the shift of attention towards the lack of glossiness in the back scattering on the hair. Noteworthy, in this case, is that there are fewer fixations on the forehead. This direct comparison shows that difficulties in the process of categorization can cause higher dwell time at the expense of other external features. In the considered case, there is a significant effect between fixations on the hair and the forehead. A one-way ANOVA between the CGI representation and the photo shows a significant effect on the hair, $F(1,104) = 6.134$, $p = .015$, and the forehead, $F(1,104) = 5.929$, $p = .026$.

The direct comparison of a real person and his or her artificial double in the same image (three geminoids and one CG character) on the final four stimuli showed that higher attention was paid to the part of the image that displayed a non-human figure. The images were placed to the left and to the right, with the boundary in the middle. The subject had to choose the non-real characters using the keyboard cursors. In this task, the images of real humans got significantly lower attention rates than their artificial counterparts. Dwell time, $t(212) = 5.171$, $p < .001$, and average fixations $t(212) = 5.792$, $p < .001$, were significantly higher than that on humans, which indicates that this particular categorization task led to higher attention on the side of the image where a non-human was depicted (the other side showed an image of a real human). The average response time was also higher ($M = 3855.47$, $SD = 2619.32$) compared to the previous single face trials ($M = 2717.89$, $SD = 2040.52$), $t(267,777) = 5.924$, $p < .001$. The average rate of correct answers was higher (83.01%) than in the other trials (60.49%), without a significant effect, $t(18) = 1.743$, $p = .098$.

3.4 Discussion

We conducted an eye-tracking study with 53 participants, to investigate if there are changes in gaze behavior of human observers toward artificial or dead characters. First, we found that the visual search of eye regions is an essential part of gaze behavior and discriminating between real and artificial humans. The analyses of the gaze sequences show that faces in the major part of the trials have not been discriminated, as long as the eyes were not considered. This was also the case using faces that differ from the human norm. Our results show that the recognition occurred between 800–1200 ms after the appearance of a stimulus and at 400–800 ms after the first eye contact. A large number of subjects only decided when their attention went to the eyes of a figure. Therefore, eye gaze behavior is similar while decision-making and in judgments of whether figures are artificial or real. Correlation analyses show that the more realistic a figure is, the longer participants need to identify it. We found further correlations between subjective ratings of faces and the duration of gaze fixations on the facial area of the eyes.

Brenton *et al.* [31] suggest that perceptual cues indicating "falsehood" are especially potent in the eyes of faces and should be investigated in context to the uncanny valley. Shorter fixation times on the eye regions as showed in this study indicate reduced eye contact and a negative feeling or less affinity. Shorter fixations were also found in simulated renderings in the experiment by Steckenfinger and Ghazanfar [297] in gaze behavior of monkeys. This study has previously not been conducted with human subjects, however, we found similar results in human depictions at intermediate levels of realism and conclude that this behavior is potentially caused by an evolutionary phenomenon in primates (RQ1). How this effect is related to cognitive processes is unknown. We assume that these aspects should be considered in the morbidity or aesthetics hypotheses in further uncanny valley research. We recommend that acceptance and ways of interaction with human-like artificial figures could be particularly improved by more credible eyes and eye-related areas.

It is important to note that behavior of the first fixations does not significantly differ from normal and task-driven gaze behavior. An interesting hypothesis arises: gaze behavior does not change due to the task because it *continuously* runs. We assume that usual human shape is continuously assessed by the perception process that can discriminate faces. We postulate that facial recognition processing includes an automatic verification of humanoids. This recognition process combines human representation with a mental illusion of a personality or helps identify familiar people from our fund of experience. The feeling of eeriness potentially awakes when facial patterns and our expectations or preferences do not match. Apparently, this kind of process runs parallel to the recognition pattern which is only activated for objects when something is classified as autonomous.

In human evolution, vision was specialized to detect and avoid threats and abnormalities. For mammals, fast movements and body language play an important role as they help to detect threats from a large distance and to distinguish friends and foes. This mechanism is particularly important in humans and in facial perception or recognition [26, 72, 169, 224]. The visual system in human improved its sensitivity to faces, which is evident in the process of perceptual narrowing [39, 187, 210]. The challenge in recognizing human characters is to differentiate threats, ambiguities, and abnormalities and is likely part of the explanation of the uncanny valley (RQ1). This study shows that we do not reject artificial figures just because they are artificial; humans reject them if they appear to be beyond our assessments – and we also reject real people when they strongly deviate from the human norm.

As our results show, strong violations and deviations from the human norm are perceived as particularly negative. However, our results are based on deduction and only show the responses of a relatively small group of participants to certain stimuli. The questions remain what defines the human norm and which deviations can be tolerated or not. What defines the dimension of human-likeness and how do we know which features on human faces are perceived as pleasant or unpleasant? To get to the bottom of these, Chapter 4 investigates which features in faces elicit positive or negative associations.

4

Preferred Virtual Faces

Based on prior findings researchers can only deduce which factors avoid eerie effects or which factors are perceived as comfortable. One way to directly learn more about the preferred preferences of users towards virtual faces is an avatar creation system. We developed the web-based *faceMaker* avatar generator and investigate how people avoid the uncanny valley for virtual faces. We determine features and facial characteristics of virtual faces which are positively or negatively perceived. Our results significantly contribute to our understanding of how attractiveness and aesthetics are able to conceal uncanny effects in virtual faces.

This chapter is based on the following publications:

V. Schwind, K. Wolf, N. Henze, and O. Korn. "Determining the Characteristics of Preferred Virtual Faces Using an Avatar Generator." In: *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY '15. London, United Kingdom: ACM, 2015, pp. 221–230. ISBN: 978-1-4503-3466-2. DOI: 10.1145/2793107.2793116

V. Schwind, K. Wolf, and N. Henze. "FaceMaker—A Procedural Face Generator to Foster Character Design Research." In: *Game Dynamics*. Cham: Springer International Publishing, 2017, pp. 95–113. ISBN: 978-3-319-53087-1. DOI: 10.1007/978-3-319-53088-8_6

4.1 Avatar Creation as a Research Method

Researchers examine the uncanny valley primarily using photos, videos of robots, or computer animations. Facial preferences are known mainly from studies of attractiveness research, which usually use photographs or processed stimuli. Given that people prefer credible and human-like characters, the question arises which appearance or details they would choose if they had a free design choice. Due to negative emotional reactions towards very realistic human figures, we assume that avatars with unpleasant or uncanny associations are given realistic but abnormal characteristics. We also suppose that positively rated faces get attractive features. However, it is still unclear which and how strong these features are, to what degree a face is stylized, and what guidelines or tools would help to avoid or prevent uncanny effects (RQ4).

An important feature of current online or role-playing games could give insights into these questions: customization of game characters using avatar generators. With a variety of parameters, users are able to personalize the appearance of virtual faces. This kind of personalization can be used to investigate how people control the appearance and which facial settings they prefer. However, such a study has not yet been conducted. This might be due to the fact that avatar creation modules in games generally allow copious adjustments, but they are limited by in-game user interfaces, based on predefined models in certain styles, or include elements which are unsuitable for research purposes.

In this chapter, we present the results of an online study using a novel web-based avatar generator called *faceMaker*. Participants were asked for their preferred parametric changes of six different stereotypes eliciting positive and negative associations (*e.g.*, heroes and villains). Thus, we investigate whether users either create realistic characters with parameters that lead into the uncanny valley, or whether they search for alternate designs. Contributions of this work are:

- Identification of important and preferred parameters of virtual faces
- Concepts of stereotypical, preferred and non-preferred virtual faces
- Design principles to avoid uncanny effects for games and interactive applications

Findings of this research can also be very helpful for developers as well as designers to create more credible and accepted game characters and could improve new avatar creation systems. Prior work on avatar creation concentrates on the examination of preferences, exploration, content, and identification of players.

4.1.1 Exploration and Case Studies

Exploration and case studies also report the outcome in games and decision making processes of game developers. Nowak and Rauh [221] evaluated avatars in terms of androgyny, credibility, homophily, attraction, and likeability. The study indicates that people are more likely to perceive an avatar as credible when it is human (instead of an object or animal). Anthropomorphism and lower androgyny of avatars increase attractiveness, credibility, and homophily. Larsson and Nerén [162] explored certain attitudes towards sexualized appearances of female stereotypes in games: females and males preferred exaggerated sexual bodies for their personal avatar over normal bodies. Chung *et al.* [48] suggested that this creation could lead to a stronger sense of self-presence and more identification to the created character. The presented studies were conducted using image renderings or pre-defined models of characters. In their study, participants were not allowed to customize the avatar faces.

4.1.2 Content

Content and design decisions of game developers are also related and have been investigated by researchers. A content analysis of 489 game characters in 60 top-selling video games conducted by Downs and Smith in 2009 [64] shows that players had more opportunities to select non-human characters (robots or anthropomorphized beings) than female characters. The study also determines that females are shown more nude, with an unrealistic body shape and sexualized clothing. A case study conducted by John [140] delivers insights of design decisions (and conflicts) between graphic designers of a game company: While the male character has to be the main figure and a photo-realistic “average guy” with anatomically correct proportions, the only female character was “designed explicitly not realistic to adjust her to heterosexual male fantasies" [140]. In this chapter, we investigate whether people who acts as designers themselves make similar decisions as the game developers.

4.1.3 Identification

Researchers also investigate the process of avatar creation itself: Heeter [114] observed teens in designing game characters and describes how girls use very high level of avatar customizations while boys rather used pre-defined characters. An exploratory study by Hussain *et al.* [130] found that 57% of online gamers decide to play a different gendered character. A study by Rice *et al.* [246] indicates similar interests in customization of an avatar for different age groups. Older players prefer higher attractiveness and show a higher homophily than younger people. Ducheneaut *et al.* [65] compare avatar creation systems of three different virtual environments. Their study shows that users emphasize the capability to change body shapes, hair styles, and hair colors. Thus, we assume that hair is an important feature in perceiving a face. It remains unclear which hair color is a significant predictor of a positive perception of a virtual face.

Summarizing research in avatar customization shows that this can be a promising approach to get more knowledge about facial preferences of users towards avatars and virtual characters. We also summarize that using avatar customization also help to understand how people would avoid the uncanny valley.

4.2 The Avatar Creation System *faceMaker*

In the following, we will describe our sytem to determine the preferred and not preferred characteristics of virtual faces. This section begins with a requirement analysis and the development of the average faces. We identify and classify the parameters of the creation system and introduce the procedure of the study.

4.2.1 Requirement Analysis

To determine facial preferences, we decided to use an avatar generation system that both allows as well as records interactive customizations of a young adult human face. Body, clothes, very old or young ages, hair styles (including skinheads), scars, tattoos, and equipments rely on the context of a game or the environment and were not included in our examination of facial preferences.

For reliable investigations of the face the following prerequisites must be met: (1) A neutral human with a minimum of characteristics, that could be preferred or rejected, (2) uniform and balanced mesh distance in order to consistently deform facial changes, and (3) must originates from the population and mean age of the surveyed target group (European or Western-oriented countries). The system to change the human face model should meet the following criteria: (1) Reach a large number of participants as possible under similar conditions as gamers usually build avatars, (2) contemporary and interactive rendering engine, (3) should be controlled by a simple, unified user interface.

Only one available avatar creation system which allow collecting user data was found [5]. Nevertheless, the core objectives of that system do not match with our requirements, because the system is based on a whole human body model and does not allow facial parametric changes of the human average face. In order to realize a contemporary interactive avatar generator that meet all requirements, we develop a new application called *faceMaker*, which controls facial changes using parametric values. According to our research design we do not directly investigate which parameters participants select in the mean, but how they build a face that they like and what they do to realize this. The application design prevents the calculation of mean characteristics in the final evaluation for two reasons: (1) Multiple features (*e.g. eyes color*) on a single scale allow no reliable assumption about a average. (2) Participants sometimes will not change the default value. Though, in order to understand which concepts participants have of a face, additional measures were taken: (1) objectives: participants should implement predefined or arbitrary concepts, (2) parameter changes, (3) areas of interests (4) assessments: participants evaluate their success in objective completion as well as the likeability, attractiveness, and gender affiliation of a face.

4.2.2 The Average Face

In order to meet all requirements of an editable human face, we decided to create a 3D model based on the average Caucasian face. Humans of that population were investigated in our target group and also regularly appear in modern video games. Furthermore, the average Caucasian face is quite well known from foregoing anthropometric or attractiveness research using digitally processed faces [102]. Due to the lack of any individual characteristics of an average face, participants can blend facial changes unhampered and need not compensate for unwanted or unfamiliar tendencies. However, previous generated images of average faces emerge from certain research questions and do not clearly refer to a large and reproducible group of faces from our target population. Due to gender related differences between average female and male faces, we decided to develop both separately.

4.2.3 Parameter Identification and Classification

To develop an avatar creation system that both evaluates general issues as well as the importance of facial features, sets of 5 general and 32 facial parameters were identified.

General Parameters

Two general questions inevitably emerge because of using the average face, and three further general questions emerge from the related work:

1. Participants should change gender related facial differences with one parameter. As previously mentioned, people are more likely to lower androgyny [221], but there is little knowledge about the degree of masculinity and femininity of virtual faces. To investigate the balance between facial stereotypes, we introduce a continuous *face gender* morphing between a female and male model, starting from an androgynous center.
2. Further studies of the average face discovered that it is more attractive than the individual faces from which it is composed [102, 159]. Skin irregularities and facial asymmetry disappear when overlapping faces, which is associated with physical attractiveness caused by mate choice in societies with higher parasite loads [92]. Nevertheless, this effect in image-processing is contrary to the assumption that more realistic or detailed characters should be used in games and have an influence on the perception of our

parameter	−			values			+	type
face gender	female			androgynous*			male	t m
face style	realistic*						cartoon	m
face details	none			half*			full	t
skin color	black			average*			white	t
hair color	black	brunette		med.blonde*	red		bright blonde	t
eyes color	black	brown*	amber	blue	lt. blue	green	grey	t
eyes shape	droopy	down	round	oval*	almond	up	asian	m
eyes opening	narrow			average*			wide	m
eyes size	small			average*			big	m
eyes height	up			average*			down	m
eyes distance	narrow			average*			wide	m
eyes orbit	bulgy			average*			covernous	m
eyes rotation	in			average*			out	m
eyebrows color	black	brunette		med.blonde*	red		bright blonde	t
eyebrows shape	pointedstraight			average*			roundhooked	m
eyebrows strength	thin			average*			thick	t
nose shape	snub			average*			hooked	m
nose length	short			average*			long	m
nose width	thin			average*			thick	m
nose bridge	thin			average*			thick	m
nose cartilage	round			average*			flat	m
forehead size	down			average*			up	m
ear size	small			average*			big	m
throat size	thin			average*			thick	m
jaw shape	triangle			average*			squared	m
jaw length	long			average*			short	m
chin shape	pointed			average*			cleft	m
cheeks shape	full			average*			scraggy	m
lips volume	thin			average*			full	m
lips size ratio	upper lip			average*			lower lip	m
mouth shape	down			average*			up	m
mouth width	wide			average*			narrow	m
mouth height	up			average*			down	m
mouth depth	backwards			average*			forwards	m
make-up eye shadow	none*						full	t
make-up lipstick	none*						full	t
make-up rouge	none*						full	t

* default value, t = texture blending, m = mesh morphing

Table 4.1: Identified and implemented parameter scales of *faceMaker* (from top to bottom): common parameters, eyes, eyebrows, nose, mouth, outer face, cheeks/jaw, and make-up.

model. Thus, we investigate whether and when the concept of a human face is consciously preferred, and introduce the investigation of asymmetric *face details* on the skin using a simple multiplied blending of a realistic photo-texture above the regular skin map of the average face.

3. Characters from the uncanny valley are often associated with the dead or zombies [316]. These are often depicted with relatively bright skin. In order to control the preferred skin type we add the *skin color* parameter, which linearly blends between dark and bright human skin, starting at the original map texture of the average skin tone.

4. To investigate the research question of whether realism or stylization is preferred in avatar creation, we developed a representative cartoon model of the average face with exaggerated and unrealistic but commonly used facial proportions. In order to implement a common way of stylization, two modeling artists aligned the model with styles of current animated family movies and comic-like game titles. The morph state was controlled by the *face style* parameter, starting from the average face without stylization.

5. The related work emphasizes the importance of hair. To investigate if hair color can positively or negatively influence the perception of a virtual face, we introduce the *hair color* parameter, which enables the user to blend among natural hair colors from black to bright blonde, starting from the average medium blonde. Certain hair styles were not included in our investigation.

Facial Parameters

In addition to the general parameters, we classified the facial sections: eyes, eyebrows, nose, outer face, head shape, mouth, and make-up. Nine state-of-the-art RPGs were examined to get an impression of prevalently used facial parameters: *Mass Effect*, *The Elder Scrolls: Oblivion & Skyrim*, *The Sims II & III*, *World of Warcraft*, *Destiny*, *Dragon Age 1 & 2*. Similar parameters were merged, obviously irrelevant (*e.g.* elf ears) were not taken into account. For usability and analysis issues, we declare a maximum number of 32. Table 4.1 lists all implemented parameters, their impact ($-/+$), and default value ($*$).

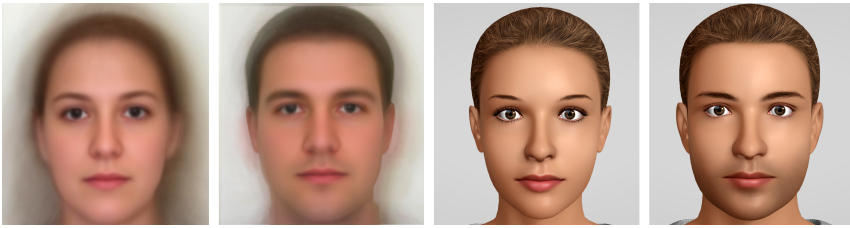

Figure 4.1: Construction of the average face. *(l)* overlay compositions of the female and male average face, *(r)* final models in 3D with skin texture (without details layer).

4.2.4 Modeling of Face and Morphings

Face repositories of 117 male and 151 female Caucasian people from 18–40 years were downloaded from the online Face Database[1] of the Park Aging Mind Laboratory (PAL) [207] and 3d.sk[2]. Both average faces (see Figure 4.1) were generated using the average face prototyping function of PsychoMorph[3], developed by Tiddeman *et al.* [312]. These templates were used as inputs of the PhotoFit feature to generate 3D models with the FaceGen[4] software. Some features of the generated 3D models were not completely useful for our purposes. Thus, both meshes were retopologized and retextured using Autodesk Maya 2014 and Mudbox 2014 by two experienced 3D modeling and texturing artists. The results of the template compositions of the processed averaged faces and the final 3D models are depicted in Figure 4.1.

Deformations of the parametrized human faces were realized using vertex displacements also know as morphings or blend shapes. Figure 4.2 shows the eight facial domains, resulting from the classification in Table 4.1 and were used as vertex selections sets while modeling as well as for explanation in the help windows of the final application. Morphings were modeled by hand.

Physical attractiveness based on facial symmetry and golden ratio is considered as a result of averaging faces [102, 160]. In order to determine whether the

[1] PAL Face Database: `http://agingmind.utdallas.edu/facedb`
[2] 3D.SK: `http://www.3d.sk`
[3] PsychoMorph: `http://users.aber.ac.uk/bpt/jpsychomorph`
[4] FaceGen: `http://www.facegen.com`

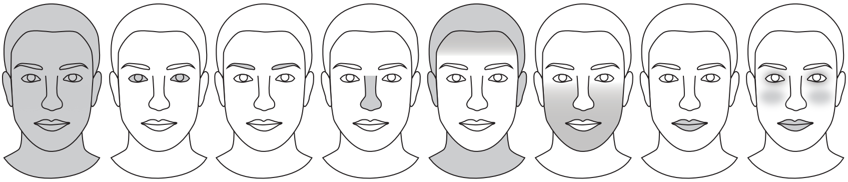

Figure 4.2: Facial domains of the *faceMaker* application. The domains were used for user-interface grouping, vertex selection of the moprhing targets, and info graphics within the final application (from left to right): common, eyes, eyebrows, nose, mouth, outer face, cheeks/jaw, and make-up.

generated average faces met the assumption of facial symmetry and golden ratio, we applied Stephen Marquardt's ϕ-Mask as suggested by Prokopakis *et al.* [237]. The mask was developed to determine physical attractiveness and to determine deviations of facial symmetry. Thus, we assume that the ϕ-Mask can be applied to the generated 3D average faces. Figure 2 shows how Marquardt's ϕ-Mask is applicable to the female as well as to the male 3D average face.

4.2.5 Objectives & Measurements

Based on our research question of how positively or negatively associated faces, we introduce 6 objectives as independent variables: (A) An personal arbitrarily avatar face. (B) An uncanny, repulsive face. The stereotypical, positive related face of an attractive (C) heroine and (D) hero. The stereotypical face of a (E) female villain and (F) male one, which is also related to negative associations. Using these objectives clearly separates facial changes attributed to positively and negatively associated faces as well as to male and females. This design also enables an evaluation of stereotypical faces which look similar to an arbitrarily designed avatar.

Parameter values, their usage (clicks), and mouse moves above the interface were dependent measures. Mouse movements were used to rotate and to slightly zoom the 3D camera to the corresponding region of the parameter. There were three reasons: (1) we make sure that possible changes on the face were noticed immediately, (2) all participants look at the same view during parameter changes, and (3) we learn more about areas of the face drawing increased attention. Furthermore, the mean displacement of all vertices and texture blendings starting from the androgynous center between the male and female average face mesh were calculated.

After generating a face, the participant had to evaluate the success in completing the objective on a 7-point Likert scale (very successful (7)/very unsuccessful (1)), likeability (very likeable (7)/very unlikeable (1)), attractiveness (very attractive (7)/very repulsive (1)), and gender affiliation (very masculine (7)/very feminine (1)). The question of success in completing the objective acts both as control function to ensure that a user complete the task correctly as well as an indicator for possible insecurities in face generation.

4.2.6 Apparatus

The online WebGL application *faceMaker* was developed using the Javascript library Three.js[1], PHP, and MySQL. Physical based skin rendering using the Beckmann shader was applied on the face model. Three directional lights and a slight ambient light (all neutral white) were used for lighting. The key light casts shadow maps. The camera can be rotated within a half circle in front of the face. The face consists of the head model, eyes, and eye lashes. For a neutral transition from neck to the décolletage, a gray T-shirt was added. The view was in fullscreen of the browser window (see Figure 4.3) with a neutral, dark grayed background.

All facial parameters were linearly controlled using horizontal sliders. Every slider has a width of 200 pixels and was labeled with the facial parameter at its left. At the right a reset button was placed to return to the default value. All parameters were gray, color sliders were equipped with a colored scale. For every new user, all sliders within a domain group as well as the groups on the interface were shuffled an distributed on the left and right column (see Figure 4.3). If a slider possessed multiple values, the current value name was displayed using tooltips below the slider.

[1] ThreeJS WebGL Engine: `http://threejs.org`

Figure 4.3: Screenshot of the final *faceMaker* application: 37 slider were organized in 8 randomly arranged containers. Default values begins between the average male and female face as well as 50% skin details.

4.2.7 Procedure

The procedure was devided into five steps: (1) Every session starts with demographic questions about gender, age, origin, and in consummating games and movies. Terms of use have to be accepted. (2) Objectives and instructions appear in a window, which have to be closed to edit the average face using all 37 sliders. Reset buttons could restore the default value. (3) Before submitting, a participant had to fulfill the four self-assessments about the created face. (4) After that, the application goes back to step 2 and was repeated until all six faces were completed. (5) If all six objective were finished, the user could view and download renderings of the submitted faces. Each objective could be processed only once. Repeats were prevented using cookies. The order was determine using the objective with the fewest participants and then successively processed in a session using a 6×6 balanced Latin square. A participant could continue the study up to 7 days, in this case all timers were stopped when leaving.

4.3 Study I: Preferred Characteristics

The first online study was conducted using a foregoing questionnaire and the *faceMaker* application. The testing phase lasted for six weeks with four beta testers. The study phase lasted three weeks. In this time, a developer monitored the application and answered to technical questions and compatibility issues via e-mail. No changes of the application were made during the main study.

4.3.1 Participants

We recruited 431 participants from 7 to 75 years ($M = 29.52$, $SD = 11.42$) using mailing lists, social networks, and advertisements. One hundred participants (35.7%) declared that they play video games more than once a week, 75 (17.9%) play more than once a month, 114 (27.1%) play infrequently, 81 (19.3%) say that they do not play video games. 204 participants (48.6%) specified that they watch movies or series more than once a week, 173 (41.2%) watch them more than once a month, 38 (9%) watch infrequently, 5 (1.2%) never. 267 participants (63.6%) have their roots in Germany, 48 (11.4%) from Poland, 14 (3.3%) from Italy, 10 (2.4%) from the U.S., and 72 (17.1%) from other countries in the western-oriented world.

1403 faces were received. 62 faces were dismissed due to following reasons: cookies were deleted during session (4), no parameter changes (32), creation time of a single face was more than 6 hours (4), participants with incompatible OS language (22). Finally, 1341 valid faces were generated by 420 participants (204 males, 210 females, 6 other/not specified) were evaluated. 145 (34.5%) participants complete all 6 objectives ($M = 3.19$, $SD = 2.19$).

4.3.2 Self-Assessments and Common Parameters

To examine facial parameters which influence the subjective perception of a face we use a regression analysis, as it is used in previous analyses of physical attractiveness (*cf.* Cunningham *et al.* [55]). How well facial preferences can depict likeability, attractiveness, and gender affiliation in relation to the average face, three multiple linear regression analyses were conducted using the enter method.

predictors	likeability β	attractiveness β	gender aff. β
face gender	-.166**	-.169**	.615** †
face style	-.036	-.102**	-.062**
face details	-.249**†	-.242** †	.075**
skin color	-.167**	-.183** †	-.023
hair color	.001	-.026	-.006
eye color	-.025	.018	.003
eye shape	.048*	.077**	-.034
eye opening	.064*	-.014	-.058*
eye size	.025	.051*	.011
eye height	.005	-.002	.000
eye distance	-.006	.008	-.025
eye depth	-.009	.016	-.017
eye rotation	.034	-.006	.044*
eyebrows color	.062*	.072*	.004
eyebrows shape	-.078**	-.041	.002
eyebrows strength	-.002	-.006	.052*
nose shape	-.106**	-.161**	.030
nose length	-.018	.009	.004
nose width	-.017	-.012	.041
nose bridge	.033	.006	.003
nose cartilage	.003	.030	-.034
forehead size	-.009	.023	-.008
ear size	-.061*	-.105**	.030
throat size	-.013	.013	.093**
jaw shape	.048	.016	.009
jaw length	-.019	-.065*	.005
chin shape	.015	.030	.031
cheeks shape	-.034	.001	.041*
lips volume	.131**†	.085**	-.073**
lips size ratio	-.014	-.021	.011
mouth shape	.138**	.086**	-.006
mouth width	.066*	.065*	.002
mouth height	-.041	-.037	.018
mouth depth	.001	.059*	-.010
make-up eye shadow	-.106**	-.044	.021
make-up lipstick	-.161**	-.131**	-.001
make-up rouge	.016	.006	-.036

* $p < .05$, ** $p < .001$, † partial correlation $r_p < -.2, r_p > .2$

Table 4.2: Standardized β-coefficients of the regression analyzes show the influences of likeability, attractiveness, and gender affiliation.

The regression equations with the facial parameters as predictors were significant for likeability, $R^2 = .463$, $R^2_{Adj.} = .447$, $SE = 1.560$, $F(37, 1303) = 30.313$, $p < .001$, $d = 2.050$, attractiveness, attractiveness, attractiveness, attractiveness, $R^2 = .456$, $R^2_{Adj.} = .441$, $SE = 1.457$, $F(37, 1303) = 29.518$, $p < .001$, $d = 2.047$, and gender affiliation, $R^2 = .630$, $R^2_{Adj.} = .620$, $SE = 1.246$, $F(37, 1303) = 60.049$, $p < .001$, $d = 2.027$. The scatterplots (not illustrated) of standardized residuals indicated that the data met the assumptions of homogeneity of variance, linearity, and homoscedasticity for all three regression analyses. No auto-correlations d were found. No effect was expected in the model of success in objective completion, which was confirmed by a fourth test, $R^2 = .059$, $R^2_{Adj.} = .032$, $SE = 1.403$, $F(37, 1303) = 1.823$, $p < .001$. We therefore assume that, there is a reliable model on the relative importance among facial parameters and final concepts of likeability, attractiveness, and gender affiliation. Table 4.2 lists all β-coefficients for each facial parameter. Related to the average face, the coefficient can be considered as a measure of the impact on a subjective impressions.

We firstly considered the results for the common parameters: Weak partial correlations among other facial parameters, excluding the parameters with the high β-coefficients (*face details*, *skin color*, and *face gender*), ranged from $-.166$ to .138 for likeability, $-.166$ to .085 for attractiveness, and $-.073$ to .93 for gender affiliation indicate that relevant judgments about the importance of the three predictors could be made.

Negative values are inversely related. The direction of that relation can be extracted from the direction of the facial parameter specified in Table 4.1. For example, both negative β-coefficients for *face gender* indicate that female faces with an increased femininity are more likeable and more attractive.

Gender affiliation and face gender correlates positively because the indicator and parameter used an increasing value for masculinity. The influence of the *face gender* parameter on the assessment of gender affiliation can clearly be attributed to the given task. The weighting is higher because participants initially changed its value according to the four gender-related objectives. Spearman's ρ reveal a strong significant correlation of *face gender* and gender affiliation, $r_s(1341) = .727$, $p < .001$. We therefore assume that the participants were able to

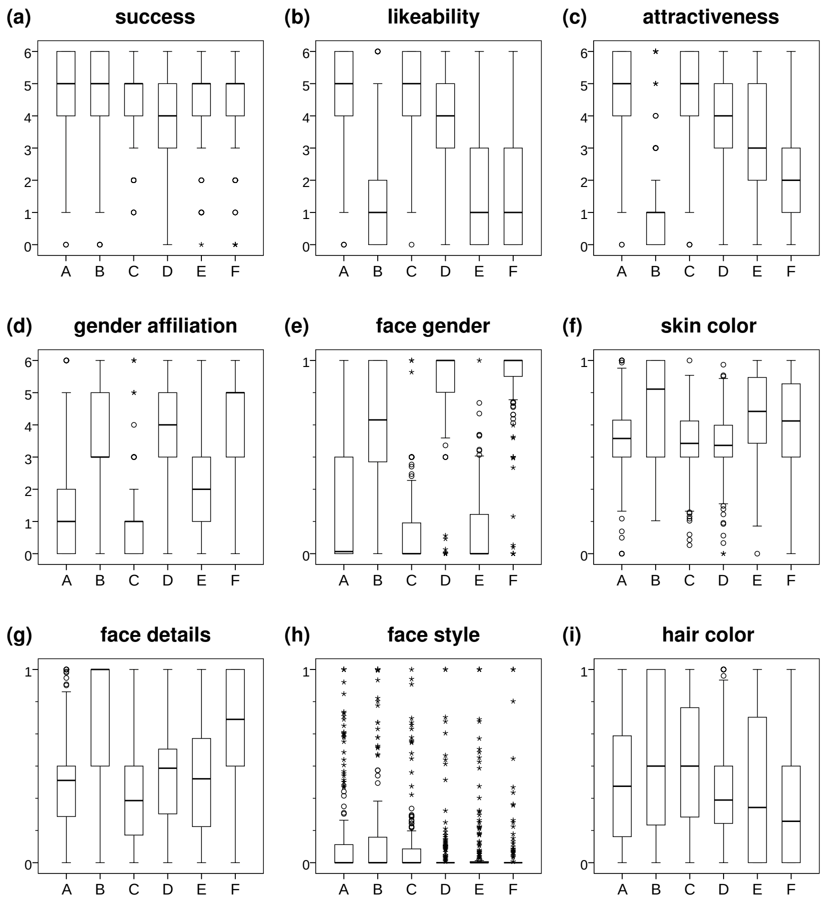

Figure 4.4: Box plots of the users' self-assessments (a-d) and actual facial parameters (f-i) for the objectives in *faceMaker*. (A) arbitrary face, (B) repulsive face, (C) female hero, (D) male hero, (E) female villain, and (F) male villain.

generate their ideas of male or female faces. However, *face details* and *skin color*, voluntarily configurable parameters without initial relations to any objective, are strongly and frequently be used to design faces more likeable (or unlikeable) and more attractive (or repulsive).

To determine possible relationships between the basic parameters *face details*, *skin color*, *face gender* and both self-assessments a further series of Spearman's rank-order correlations were conducted. A two-tailed test of significance indicated that there was a positive correlation between likeability and attractiveness, $r_s(1341) = .743$, $p < .001$, but we assume that the parameters are not entirely redundant. Inverse relationships between *face details* and likeability, $r_s(1341) = -.400$, $p < .001$, as well as between *face details* and attractiveness, $r_s(1285) = -.424$, $p < .001$, were found. Both negative correlations of *skin color* towards likeability, $r_s(1341) = -.325$, $p < .001$, as well as attractiveness, $r_s(1343) = -.300$, $p < .001$, were weaker.

The box plots from Figures 4.4b and 4.4c show the similarity of the two assessments as well as the contrary relationships among them and the parameters *face details* and *skin color* shown in Figures 4.4f and 4.4g. Due to lack of normal distribution a series of Kruskal-Wallis tests was conducted to find contrasts among these tendencies. The tests showed significant differences of the distributions of likability, $\chi^2(5, N = 1341) = 644.832$, attractiveness, $\chi^2(5, N = 1341) = 552.770$, *face details*, $\chi^2(5, N = 1341) = 320.970$, and *skin color*, $\chi^2(5, N = 1341) = 184.985$, across all objectives (with all $p < .001$). Table 4.3 lists the post-hoc rank differences of all pairwise comparison tests.

Significant rank differences were found by pairwise comparisons between the ratings of success through an independent-samples Kruskal-Wallis test, $\chi^2(5, N = 1341) = 60.365$, $p < .001$, in the category of male heroes towards arbitrary ($H = 245.4$, $p < .001$), repulsive ($H = -185.3$, $p < .001$), female hero ($H = -138.2$, $p < .001$), and female villain ($H = -141.8$, $p < .001$) faces. In spite of quite positive ratings of success (see Figure 4.4a), we therefore assume that it was relatively more difficult for users to generate faces of male heroes.

	self-assessment		facial parameter		
test	**likeability**	**attractiveness**	**face details**	**skin color**	**face gender**
A-B	576.4**	-630.0**	-442.6**	-289.0**	-335.2**
A-C	68.8	-75.8	98.3	60.8	101.0
A-D	68.1	96.2	-45.5	69.2	-525.9**
A-E	480.2**	273.9**	-29.4	-239.1**	85.3
A-F	507.0**	415.1**	-298.6**	-159.6**	-553.3**
B-C	-645.2**	-705.8**	540.8**	349.8**	436.1**
B-D	-508.3**	-533.9**	397.1**	358.2**	-190.7**
B-E	-96.3	-356.1**	413.1**	49.9	420.5**
B-F	-69.4	-214.9**	143.9*	129.4*	-218.1**
C-D	136.9*	171.9**	-143.7*	8.3	-626.8**
C-E	548.9**	349.7**	-127.7*	-299.9**	-15.6
C-F	575.8**	490.8**	-396.9**	-220.5**	-654.2**
D-E	412.0**	177.7**	16.0	-308.3**	611.1**
D-F	438.8**	318.9**	-253.2**	-228.8**	-27.4
E-F	26.8	141.2**	-269.2**	49.4	-638.6**

* p < .05, ** p < .001

Table 4.3: Rank differences (H) among all objectives: pairwise Kruskal-Wallis post-hoc tests with asymptotic significances (2-sided). (A) arbitrary face, (B) repulsive face, (C) female hero, (D) male hero, (E) female villain, and (F) male villain.

4.3.3 Facial Preferences

In the following we report results of the regression analyses related to parameters within the facial domains.

Eyes and Eyebrows

The *eye shape* is highly significant for both attractiveness ($\beta = .077$, $p = .001$) and likeability ($\beta = .048$, $p = .030$). The β-coefficients are slightly positive which means that more almond, upturned, or Asian-shaped eyes in relation to the average face are more likeable and attractive than downturned or droopy eyes. Interestingly, the eyes were slightly downturned ($\beta = .044$, $p = .019$) to shape more masculine faces. *Eye opening* controls the distance between the eye lids.

The β-coefficients for likeability ($\beta = .064$, $p = .004$) are positive and gender affiliation ($\beta = -.058$, $p = .002$) negative, which means that slightly opened eyes are perceived more attractive and less masculine. A higher β-coefficient for *eye size* ($\beta = .051$, $p = .027$) indicates that slightly increased eye size is perceived more attractive. Likeability ($\beta = .062$, $p = .016$) and attractiveness ($\beta = .072$, $p = .005$) increase for lighter eyebrows than the average. Straight and pointed *eyebrow shapes* ($\beta = -.078$, $p < .001$) lead significantly to less likeability, and stronger eyebrows ($\beta = .052$, $p = .004$) to more masculine faces.

Nose, Outer Face, and Head Shape

Only the shape parameter of the nose group significantly led to more likeability ($\beta = -.106$, $p < .001$) and attractiveness ($\beta = -.161$, $p < .001$). When viewing negatively associated faces we already noticed the often used hooked nose, while positive-associated faces were sometimes equipped with straight but mainly with snub noses. This effect can be observed for both genders. Wider throats ($\beta = .093$, $p < .001$) and scraggy cheeks ($\beta = .041$, $p = .031$) were used to generate more masculine faces. Ear sizes were reduced to generate more attractive ($\beta = -.105$, $p < .001$) or likeable ($\beta = -.061$, $p = .011$) faces. The jaw length was also reduced for more attractive faces ($\beta = -.065$, $p = .007$).

Mouth and Make-Up

A dynamic facial parameter is the volume of the lips, which was strongly increased to generate more likeable ($\beta = .131$, $p < .001$) and attractive ($\beta = .086$, $p < .001$) faces, but reduced to generate more masculine faces ($\beta = -.073$, $p < .001$). Mouths were shaped upturned ($\beta = .138$, $p < .001$; $\beta = .086$, $p < .001$) and slightly tighter ($\beta = .066$, $p = .005$; $\beta = .065$, $p = .004$) in relation to the average to generate more attractive and likeable faces. Slightly protruding lips led to more attractiveness using mouth depth ($\beta = .059$, $p = .012$). Textural blending of make-up layers has different effects on the favorability of each feature. More eye shadow ($\beta = -.106$, $p < .001$) led to lower ratings of likeability and more

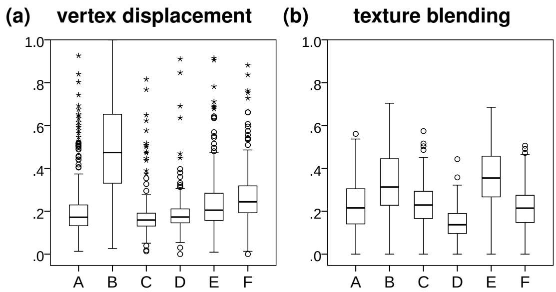

Figure 4.5: Differences of vertex displacements and textures transparencies between the objectives. An average vertex blending (a) of 0.1 is equal to 10% away from the maximal morphs. A transparency blending (b) value of 0.1 is equal to an average blending of 10%.

lipstick led to lower ratings of likeability and attractiveness ($\beta = -.161$, $p < .001$; $\beta = -.131$, $p < .001$). These features were often used to create villain faces (see Figures 4.6e and 4.6f).

4.3.4 Stereotypes and Differences from the Averages

Due to different vertex displacement between the facial morphings, it was not possible to compare the objectives and the original average faces by using the mean values of the parameters. Thus, we calculate the mean displacement of all vertices as well as texture blendings within the mesh to determine how strong a face differ from the androgynous average face (see Figure 4.5). All vertex displacements are normalized to their tenth starting from the androgynous center of both average faces. This means that a displacement of .1 is equal to a mesh blending to both the average male as well as the female face respectively.

Further independent Kruskal-Wallis tests shows significant rank differences between the vertex deformations, $\chi^2(5, N = 1341) = 446.1$, $p < .001$, and texture blendings, $\chi^2(5, N = 1341) = 368.4$, $p < .001$, of stereotypical faces. Pairwise comparisons of vertex displacements reveal adjusted significances for all test pairs excluding F-A ($p = 1$), F-C ($p = 1$), B-E ($p = .093$), and A-C ($p = .469$). Mean

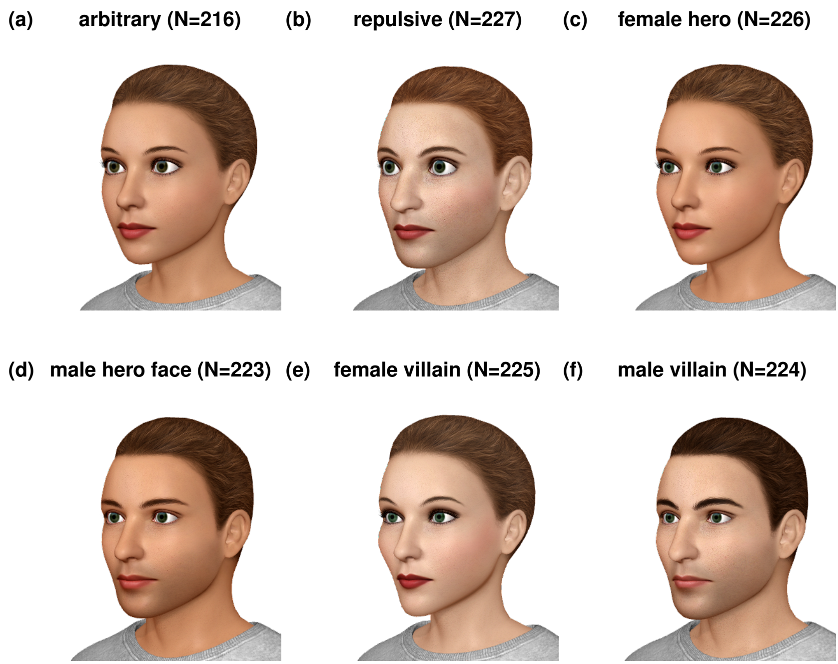

Figure 4.6: Average stereotypical faces according to the six objectives.

texture transparencies differ between all test pairs excluding D-A ($p = .495$), F-C ($p = .173$), and A-C ($p = 1$). Both tests underscore the similarity between the average arbitrary face and the heroine shown in Figures 4.6a and 4.6c.

4.3.5 Realism versus Stylization

One of our research question was whether and how participants used a stylized cartoon face and how they controlled the visibility of details if they had a choice. Thus, each participant had the chance to stylize a face (or not) and to completely remove (or add) texture details. However, in 71% of all cases the face was not morphed by the parameter *face style*. In contrast, only 11.8% of all faces received no *skin details*. In the cases when stylization was applied, participants choose a very low degree of cartoon-like proportions (see Figure 4.7).

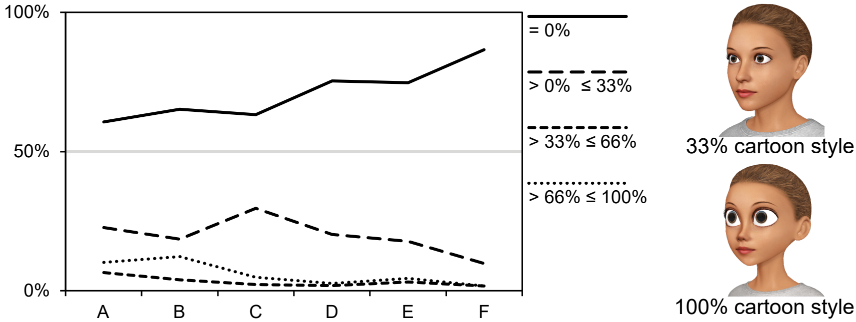

Figure 4.7: User stylizations in *faceMaker*: Heroines get most stylized morphings from a blending range of 0–33%. (A) arbitrary, (B) repulsive, (C) heroine, (D) hero, (E) female villain, and (F) male villain.

To understand why some participants choose these facial proportions, we examined differences between the objectives and discovered an increase for the heroine face between a morphing range 0% to 33.0%. This caused the face to appear almost realistic, but seeming younger through childlike proportions (see Figure 4.7) – attributes that are perceived as particularly attractive [160]. Light *skin details* give the impression of a healthy but natural and realistic skin. Also noticeable at this point is, that villains were rarely stylized but a few people also use strong stylization (from 66%–100%) to generate repulsive faces.

4.3.6 Task Completion Time

The mean task completion time (TCT) for generating a face was about 8 minutes ($M = 8.67$, $SD = 19.09$) and slightly varies between the objectives. Participants took most time for the arbitrary face ($M = 11.15$, $SD = 27.65$), male villain ($M = 8.62$, $SD = 15.39$), and the heroine ($SD = 9.79$, $SD = 25.14$). The male hero ($M = 7.82$, $SD = 9.94$), the repulsive ($M = 7.43$, $SD = 18.18$), the female villain ($M = 7.32$, $SD = 11.49$) were below the average times. A one-way ANOVA showed no significant differences of the TCTs between the objectives, $F(5, 1335) = 1.396$, $p = .223$.

4.3.7 Mouse Clicks and Moves in *faceMaker*

The average count of a single slider change per user is $M = 1.438$ ($SD = 1.438$) and of mouse movements is $M = 4.069$ ($SD = 3.714$). Figure 4.8 shows a sorted list of both measurements. Comparisons with predictions from the linear regression analyses show that some important values were also frequently used (*e.g.*, *face gender*, *face details*, and *skin color*). This does not apply to all values. A particularly unreliable predictor for likability or attractiveness is *hair color*; however, mouse activities indicate, that the hair color is quite important in designing avatars. *Eyes color*, *eyes depth*, *eyes shadow*, and *eyebrows strength* as well as the head shaping parameters *cheeks shape*, *nose shape*, and *face style* were no reliable predictors for likeability and attractiveness but were frequently manipulated.

4.3.8 Discussion

This study puts the negative effects of the uncanny valley and facial properties on virtual human faces into context. The basis of this investigation is the average face. Our results indicate that smooth skin, natural skin color, and human proportions are the most relevant factors to avoid negative feelings. Users rarely choose a virtual face with an obvious cartoon-like look. This led to new findings in the uncanny valley and avatar creation research: If having a choice, people were more willing to create human-like instead of cartoon-like faces known from animation movies or video games. So people consciously move into the direction of familiar human-likeness and “bridge” the uncanny valley using very attractive features for faces they like.

We show which facial parameter influence attractiveness or likeability. Negatively associated faces get features that deviate from the human norm. Villains, for example, get strong make-up, and frequently are given striking hooked noses (*cf.* Figure 4.6e). Female villains get distinctive cheek bones and lip stick. Male villains have straight eyebrows and distinctive jaws. Repulsive faces are exaggerated with unnatural violations against the human average. Negative concepts of human faces generally get unnatural bright skin (associated with an uncanny zombie-look) and a very strong overlay of realistic skin details.

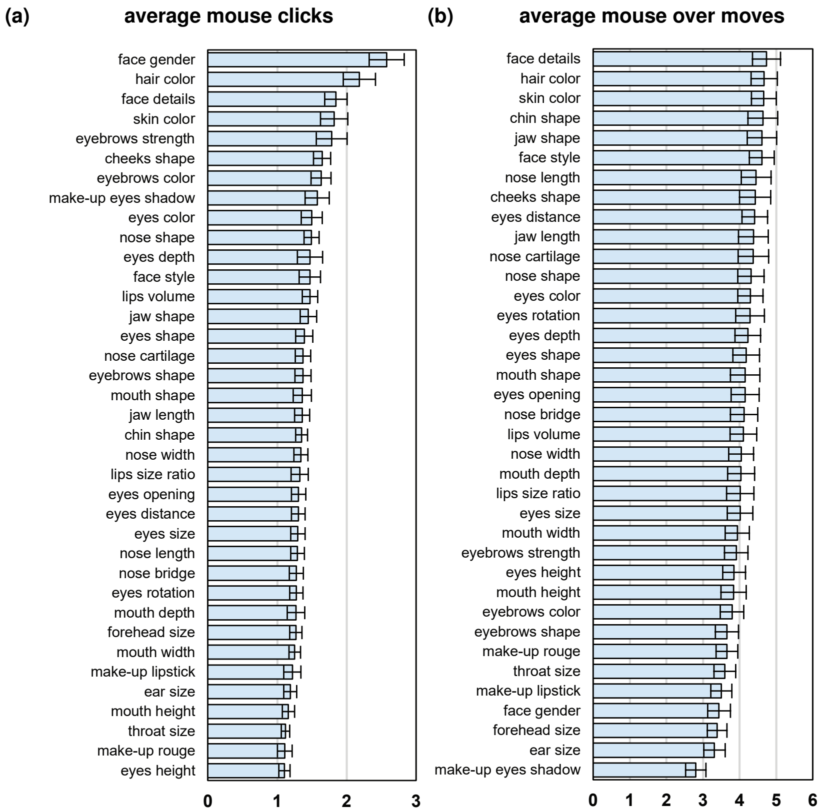

Figure 4.8: Average clicks and mouse movements per item on every facial parameter. Clicks (a) indicate changes; mouse moves (b) were used to rotate and to zoom the camera to the corresponding area. Error bars show the 95% confidence interval.

In contrast, positively associated faces get smooth skin, realistic proportions, and natural average skin color (of the surveyed population). We also notice that female faces are also equipped with full lips, snub nose, and slightly upturned eyes. Males get strong eyebrows, downturned eyes, a larger throat, and thin lips. Thus, some previously mentioned guidelines (smooth skin, less photo-realistic textures) in the related work to avoid the uncanny valley were confirmed [192] and extended. Interestingly, participants chose for the skin realism of appealing faces a skin texture in a golden ratio of ca. 38% for the photo-texture and ca. 62% for the perfect single colored skin beneath. This supports the assumption of previous work saying that slight imperfections in faces are still more attractive than perfect faces [136, 304].

Furthermore, stylization is rarely used and, if at all, only with a low value on heroines, which gives them a more youthful look. Some repulsive faces get stylization in combination with strong skin details in order to depart from the human average. We therefore conclude that participants consciously do not violate against human norms but rather search for both credibility and attractiveness to avoid negative associations. Positively associated figures are closer to the average face than negatively rated faces (*e.g.* repulsive faces). This is confirmed by the calculation of average vertex displacements and texture blendings.

There were very small and insignificant differences between the voluntary modeled face and the female hero – also for the parameter *face gender* (*cf.* Figures 4.6a and 4.6c). About 74% of the users of *faceMaker* choose a rather feminine type when designing an arbitrary face. Due to an almost identical number of male and female participants, we assume that an attractive female face is generally preferred. Furthermore, concepts of stereotypical (hero, villain) faces are very clear and should not include any androgyny. Both facts confirmed the results of Nowak and Rauh [221].

4.4 Study II: Face Classification

In a second study of the *faceMaker* data, which we present in the following section, we use clustering analysis to understand which *kinds* of faces were created without considering the objectives directly. Cluster analyzes are well-established means to investigate data with uncertain or unknown (sub-)structures. We assume that our objectives have delivered a higher contrast between different kinds of faces and relieve clustering without using any tasks for participants, however, we are able to deeply explore these clustered data and to learn more about potential subsets of faces which have not been quantified using the linear regression analyzes in the first study.

4.4.1 Participants

For the second study using *faceMaker*, we collected the data of 569 participants (313 males, 247 females, 9 other/not specified) who created 1730 faces. The mean age was 30.58 (SD = 11.86). The participants were mainly recruited via mailing lists, social networks, and advertisements. One hundred twenty seven (22%) participants pointed out that they play games daily, 108 (19%) more than once a week, 68 (12%) once a week, 40 (7%) once a month, 125 (22%) play infrequently, 101 (18%) never. Eighty-one (14%) participants pointed out that they watch movies every day, 192 (34%) more than once a week, 171 (30%) once a week, 53 (9%) once a month, 62 (11%) watch infrequently movies, and 10 (2%) never.

4.4.2 Clustering and Multi-Dimensional Scaling

To group and visualize the procedural faces systematically, cluster analysis was employed. Among different clustering algorithms, we applied the expectation maximization (EM) algorithm for several reasons: EM does not require a pre-defined number of cluster (as kMeans); the method iteratively searches for the maximum likelihood, and was adopted in the previous work, *e.g.* for face detection as used by Rujirakul *et al.* [258]. It is known that different cluster algorithms often produce inconsistent results. Different distance measurements can be used to validate the consistency of an algorithm. We decided to use the Euclidean distance metric because other metrics such as Ward's method tend to establish equal cluster sizes. To understand how the clusters are related to each other, we developed a spatial map to visualize the results. The distance matrices were used to conduct a multi-dimensional scaling (MDS). This approach cannot provide the accuracy of a face classification or the complete clustering, however it vividly illustrates the similarities or differences of faces. EM clustering was conducted in Weka, distance metrics, and multidimensional scaling was computed using R cmdscale[1].

4.4.3 Analysis and Statistical Results

Based on the training set only given by the created facial parameters, EM clustering delivers 6 nodes: Cluster 0: 282 (17%), Cluster 1: 362 (21%), Cluster 2: 169 (10%), Cluster 3: 209 (12%), Cluster 4: 229 (13%), Cluster 5: 457 (27%). The 6 objectives that people created were verified using a cluster assignment. Cluster 0 was assigned to F (male villain face), Cluster 1 to C (attractive heroine face), Cluster 2 to A (arbitrary face), Cluster 3 to B (repulsive face), Cluster 4 to E (female villain face), and Cluster 5 to D (male hero-villain). We would like to note that the results of the EM algorithm and the assignment procedure are incidentally identical to the number of objectives. The parameter results of the cluster analysis were added to the plotted diagram of the multi-dimensional scaling (see Figure 5). 55.0% of the faces were incorrectly clustered instances. The class attribution table (not illustrated) reveals that Cluster 1, for example, shares 257 instances with objectives A and C, which could be explained by the tendency of participants to create faces in the arbitrary task that are similar like in the female-hero-task.

The plotted MDS map (Figure 4.9) reveals a dense filament including two main consolidations at the top and clusters of faces that people prefer to create. Parametric values of the cluster centers were used to render faces and were placed on the MDS map. Two main clusters were connected in the main filament structure. Cluster 1, 2, and 4 as well as the average faces of objective C and E were on the female "side" on the map, and Cluster 0 and 5 as well as faces of objective D and F were on the male "side" of the map. Since all parametric changes were weighted equally, the arrangement and the results of this cluster analysis reveal that the sum of parametric changes is made according to gender and appeal. The arbitrary face (A), which participants created without any restrictions, is very close to the female cluster of faces. We also see that the cluster center of female

[1] cmdscale – A R-library for multi-dimensional scaling: Retrieved June 2016 from `https://stat.ethz.ch/R-manual/R-devel/library/stats/html/cmdscale.html`

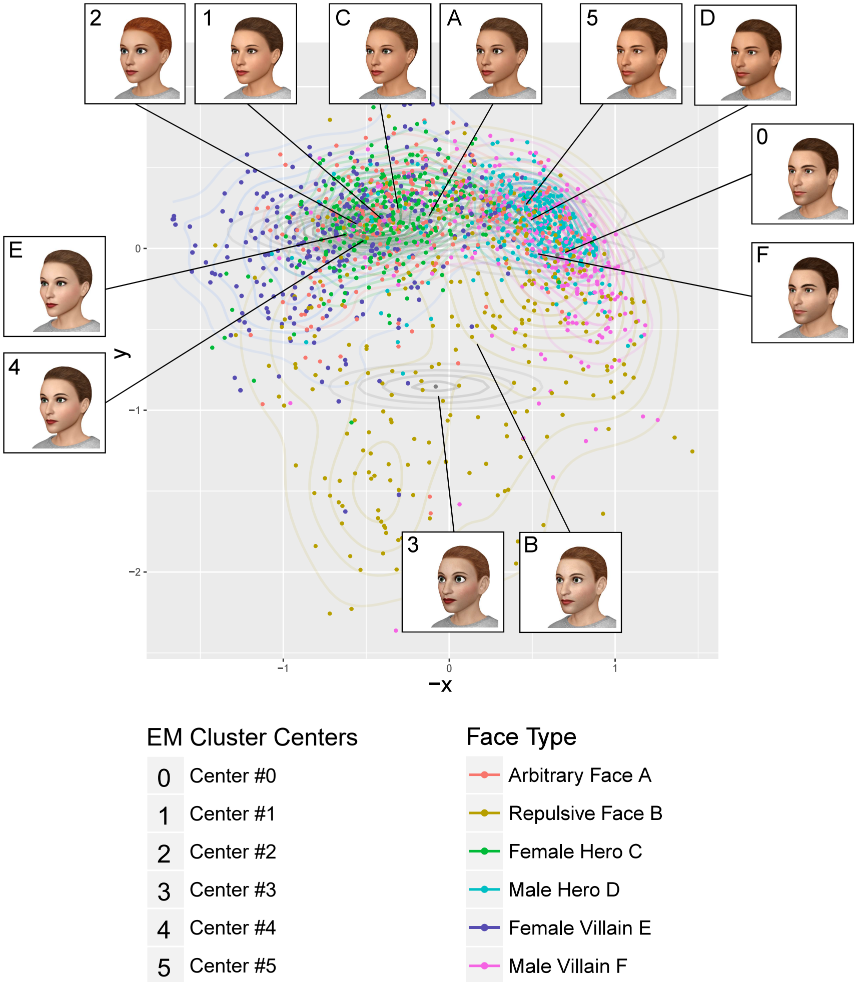

Figure 4.9: Plot of the MDS analysis of the *faceMaker* data including 1730 faces from 569 participants and location of the average faces (A-F) resulted from the users' objectives and the 6 cluster centers (1-6) delivered by expectation–maximization (EM).

villains (4) is much closer to the female heroines (1) than the male villain center (0) from the male hero (5). The cluster center of the repulsive face is outside the cluster centers of male or female faces.

The MDS map of the cluster analysis reveals no further subsets in our data but validates and strengthens the results in our first study. However, the clustering analysis provides further insights in the data structure and shows that facial properties of the created faces depend on gender and appeal. Male faces tend to be placed at positive x-values; female faces at negative ones. Appealing faces (heroes) tend to have positive values on the y-axis, unappealing (repulsive) ones have negative values. Using this approach, we are now able to derive stereotypical faces and avatars. For example, to find very female facial properties we can look at samples very close to the cluster centers 1 and 2. In contrast, very repulsive faces are outside of the main cluster of females and males. The repulsive face samples show a very distributed pattern. This can be explained by the fact that repulsive faces have no certain features or patterns; however, they deviate from the human ideal and human average proportions that people try not to violate while creating appealing faces.

4.4.4 Discussion

In the second study using *faceMaker*, we investigated the parametrized data of facial characteristics in a cluster analysis based on 1730 procedural faces created by 569 participants. The data was clustered using EM clustering and we found the same amount of cluster nodes as given objectives (arbitrary face, repulsive face, female face, male face, female villain, male villain). Through cluster alignment, we determined which cluster generally corresponded to the stereotypical average face. Plotting the results of a MDS analysis reveals how the faces are related and which face types deviate from or correspond to the human average and human ideal. The map reveals two main distributions of male and female faces and shows how participants create faces according to their concept of stereotypes. Class attribution reveals that Cluster 1 shares 257 instances with the arbitrary and female hero face. If having a choice, our participants are more willing to create female faces than male faces. Positively associated faces (female and male heroes) get smooth skin, realistic and attractive proportions, and natural average skin color.

Villains get bright skin, distinctive cheek bones, and exaggerated jaws. Female villains receive strong lip-stick; male villains get strong eyebrows. Repulsive faces get features that strongly deviate from the human norm, which is divided into two main distributions of male and female faces. They are exaggerated with unnatural violations against the human average. Thus, using these procedural types of faces, we are now able to look into preferred or not preferred stereotypical concepts of avatar faces and can deduce which facial properties are rather preferred or not preferred.

4.5 Discussion

4.5.1 Summary & Context

The results of our first study showed that participants used attractive and appealing facial characteristics to avoid uncanny effects (RQ4). Surprisingly, the participants avoided too high degrees of stylization as well as too high degrees of human realism. They put more effort into creating appealing and very beautiful faces with ideal proportions and appealing features such as smooth skin and healthy skin color. This indicates a strong aesthetic desire, possibly caused by evolutionary reasons and due to optimal mate selection. The choice of attractive and healthy facial features shows that people are more likely to avoid realism and on average do not deviate from human norms. We assume that by using healthy skin tones avoiding atypicalities from the human average our participants also reduce potential perceptual conflicts and mismatches.

In context with the discussion, if the uncanny valley is either caused by categorical perception or by perceptual mismatch, there is an interesting finding in our second study as our participants choose multiple deviations and abnormalities from the human average in order to decrease affinity and, thus, to actively increase perception tension. One interesting aspect is eye size, which is mentioned in multiple studies eliciting eerie effects [33, 145, 287]. The results from our first study indicate that the eyes were enlarged in order to make the face look younger. However, further enlarged eyes combined with other abnormal properties reduced likability. In particular, the analysis of the data of the second study shows, that multiple deviations from regular features are perceived as uncanny and unpleasant.

Figure 4.10: Random selection of 15 user generated faces via *faceMaker*.

Skin details such as pale skin tones and unnaturally enlarged eyes were combined to increase the perceived eeriness. That high skin realism elicits expectancies of high eye realism but not *vice versa* is confirmed by other researchers and explains why makeup increases (and not decreases) facial attractiveness "but a dead person's eyes do" [185, 187]. These findings further support the perceptual mismatch hypothesis, which predicts that multiple conflicting cues increase the perceptual tension predicted by Moore [210], Chattopadhyay and MacDorman [39] and negative affinity in the uncanny valley predicted by Mori [213].

This result potentially allows drawing a conclusion about the relationships of perceptual mismatch, the morbidity, and aesthetics hypotheses. Attractiveness can increase affinity, independently from the fact that a character is "good" or "evil", as social aspects are secondary for passing on the own genes. As indicated by evolutionary aesthetics in the related work (see Section 2.3.5), humans respond positively to *any* aesthetic stimuli that are conducive to survival and reproduction. This would include both human and non-human entities, however with a higher effect size at increased human-likeness due to its relevance in mate selection. Morbidity cues are similar, but present a higher threat when they occur in the own species. As our results show, people use attractive features to avoid uncanny effects and use multiple disfigurements or atypicalities to deviate from the human norm, which potentially confirms this hypothesis. The relation between morbidity and aesthetics would be the degree of human-likeness as their relevance and effect size increases when the own species is affected.

A further interesting finding is that negative associations are considered differentiated: evil characters (villains) were made much more attractive than faces that were made to be uncanny and repulsive. Why do evil characters receive fewer perceptual conflicts? The *faceMaker* results indicate that the notion of the uncanny seems to be different from characters that appear good or evil. This indicates that social aspects are independent of aesthetic preferences and from perceptual tension and the uncanny valley.

4.5.2 Design Implications

We finally recommend providing avatar customization for optimal affinity. Escaping the uncanny valley by design not only means to design according to aesthetic preferences of the users but also to escape by providing design opportunities. Avatar customization allows individual adjustments of the desired appearance for the own avatar. Furthermore, people put a lot of effort into the process of creating avatars they like to have. This process not only avoids that people getting an avatar they do not like but further increases the emotional connection and identification [65] enhancing its positive effects. These factors may decrease the uncanny valley effect, and we strongly suggest avatar creation or avatar customization systems in video games for best affinity.

Further findings from our results are interesting for game designers: Male faces get generally more details on their skin than females. Villains get more details than heroes, but less stylization. Female heroes and female villains have a slightly brighter skin than their male stereotypes, but without significant differences. The importance of parametrized hair in avatar creation was confirmed (*cf.* Ducheneaut *et al.* [65]) with measurements of the mouse activity in our first study. Hair color is no predictor and, thus, irrelevant for perceived likability or attractiveness. We assume that there are outbalanced preferences for hair of a virtual face.

4.5.3 *faceMaker* as a Research Toolkit

The avatar and face creation system faceMaker enables researchers to conduct valid online studies reaching a large number of participants. It opens up new branches of customized avatar related research instead of presenting predefined images. The program can be used as experimental apparatus in laboratories as well as web-based application in the large. This is now also possible for other researchers who investigate the human perception or attitude towards virtual avatars. For investigations of differences in other cultures or different ages, it is necessary to use average faces from different ethnic groups, for example Asian or Africans faces.

Researchers in the fields of human-computer interaction, games, social sciences, psychological sciences, medicine, and other disciplines are now able to use faceMaker for their own purposes. The application is available for free disposal, and can benefit through new findings by investigating users and their generated avatars instead of analyzing predefined face models only. Game developers can use our system to conduct pre-studies to investigate which facial properties main characters or stereotypes in their game should have. They can also use the system to optimize location-based changes or questions about their main characters or user interfaces. Furthermore, we suggest future research about how users create procedural faces they already know (as celebrities), look like themselves, or investigate gender or cultural differences in the face creation process.

Game developers can now deduce which avatar faces are preferred and which not preferred by their users. Knowledge about preferred or not preferred facial characteristics of avatars gained by faceMaker can be considered in design decisions about the appearance of game characters and future avatar generators. The usage of parameters provided in faceMaker allows optimizing such avatar generators. For example, we showed that gender, hair color, and eye shape are the most changed parameters. This knowledge could be considered in the development and optimization of user interfaces of game character generators in RPGs.

Likeability and physical attractiveness were used as indicators to learn more about features that make a virtual face more likeable or appealing. This could be extended using further indicators of interest. Shares of more realistic attributes (*e.g.*, freckles, moles, acne, wrinkles, beards, hair, more cartoon styles) affects the impression of skin, and should also be further investigated. Using *faceMaker* we plan to collect more data over a longer period of time to investigate gender related differences, cultural preferences, and possible differences between age groups. Technical limitations include the lack of volumetric hair and fine geometrical displacements instead of bump maps.

To the current date (2018-07-15), 4366 faces have been collected, which can be used for further evaluation and research purposes. Therefore, we provide the generated faces, aggregated results, and the *faceMaker* application for free disposal[1]. Due to our positive experiences, we strongly recommend the average face as base model for design decisions in game development or in virtual face related investigations. Other questions which have arisen are whether people perceive a face differently than the user who created it and how they would rate faces they already know or that look like themselves.

[1] `https://github.com/valentin-schwind/facemaker`

5

Virtual Animal Characters

In this chapter, we investigate if the uncanny valley is not only related to human-likeness. Prior research assumes that atypical features trigger unpleasant effects only in human-like entities. We found that animal realism as well as atypicalites can reduce affinity in virtual animals and evoke uncomfortable feelings. Quantitative and qualitative findings have shown similar effects for animals to those Mori predicted for human-likeness. We discuss how these findings can be integrated into existing theoretical frameworks and argue that these factors potentially belong to the same cognitive mechanism; namely the uncanny valley. Finally, we contribute with recommendations and design implications for virtual animal entities in HCI.

This chapter is based on the following publication:

V. Schwind, K. Leicht, S. Jäger, K. Wolf, and N. Henze. "Is there an Uncanny Valley of Virtual Animals? A Quantitative and Qualitative Investigation." In: *International Journal of Human-Computer Studies* 111 (2018), pp. 49 –61. ISSN: 1071-5819. DOI: `https://doi.org/10.1016/j.ijhcs.2017.11.003`

5.1 Context and Research Rationale

In previous chapters, we showed how virtual characters affects our gaze behavior and how people prevent or avoid unpleasant effects of the uncanny valley using appealing characteristics of human faces. The related work also showed some of the multiple conditions that potentially lead to the phenomenon. In order to learn more about these causes it is necessary to investigate if the phenomenon is not only related to human-likeness (RQ3). In the following, we examine if there are similar effects of the uncanny valley for non-human entities – specifically for virtual animals.

A hypothetical "uncanny valley of animals" would mean that Mori's dimension of human-likeness is not only related to humans and has to be extended, or that the phenomenon appears in a different form (or not at all). In the following, we therefore investigate whether the uncanny valley is applicable to *virtual* animals since it is assumed that the effect also occurs for virtual humans and virtual stimuli can easily be manipulated. We aim to answer the following research questions: (1) Can findings of the uncanny valley be transferred from human-like to animal-like virtual characters? (2) Which factors potentially cause unpleasant effects for virtual animals? (3) Which design implications result from these findings? Research has found that when using CG characters the uncanny valley appears between intermediate and high levels of virtual realism and using atypical entity features, which indicates there are potentially two kinds of uncanny valleys [40, 41, 43, 103, 145, 179, 192, 209, 210, 287]. We hypothesize that both aspects can exist for human observations of non-human CG characters (virtual animals) and operationalize the effect as lower familiarity ratings.

- H1: A virtual animal rendered at high levels of realism is perceived less familiar than using photo-realism or a stylized image.
- H2: At high levels of photo-realism atypical virtual animal features decrease familiarity more than at lower levels of photo-realism.

Furthermore, it is important to understand *which* factors trigger potentially unpleasant effects using virtual animals. By contrasting these findings with previous work for human-like characters we are able to extend our knowledge of existing frameworks that try to explain the uncanny valley or potentially allow the development of a new overarching framework which considers humans *and* animals.

This work is also important to learn more about when and how virtual animals should be used in virtual contexts. Due to their symbolic and allegorical character, designers, storytellers, and engineers often prefer animal characters instead of humans. Artificial animals are frequently used in entertainment and advertising as well as for therapeutic and educational purposes [148]. Beside the positive effects of using artificial animals, there are also reports of negative experiences with virtual or stuffed animals. Noteworthy examples of negative responses are critiques of the CG cat *Azrael* in the movie *The Smurfs*.

One feature of Azrael is that the cat expresses itself using human emotions and its behavior is human-like (see Figure 5.1). Burr from the Boston Globe calls that cat a "creepy animal CGI" [34]. Duralde from The Wrap, in reviewing *The Smurfs 2* movie, states that *Azrael* "[...] never feels like a real feline; turns out there's an uncanny valley for animals, too" [68]. In art and literature, reactions to stuffed and composed animals, for example through taxidermy, have also been associated with uncanny sensations (*cf.* Gutierrez [106] and Powell [235]).

In the following, we present the results of multiple studies to gain deeper insights into the human perception of virtual animals. First, we investigate the effect of different levels of realism on the perception of virtual animals and analyze qualitative feedback provided by our participants in an online survey. The second study examines the effects of stylization and anthropomorphic emotions using virtual animals. We discuss the integration of an uncanny valley of animals into existing theoretical frameworks. We conclude with design principles for avoiding the uncanny valley effect using realistic animal characters, and finally recommend to consider virtual animals for a better understanding of the uncanny valley.

Figure 5.1: CGI character *Azrael* from *The Smurfs* (2011), Source: DVD, Sony Pictures Animation, Columbia Pictures.

5.2 Research on Virtual Animals

In the following, we highlight work that shows the explicit usage and importance of animal characters in video games and HCI. Afterward, we discuss previous work on the uncanny valley related to (virtual) animals.

5.2.1 Virtual Animals in Video Games & HCI

Based on a survey of game magazines from 1988 to 2005, Miller and Summer [206] report that over the years, animals have been the main character of up to 14.6% (from 1991 to 1993) of all video games. Furthermore, animals have been the most prominent enemy in up to 36% (from 1994 to 1996) of all surveyed games. Virtual animals are also used as companions in educational technology. Chen *et al.* [44], for example, purposely selected in their educational game an animal as the main character whose traits and behaviors are governed by the student's learning profile. Similarly, Hswen *et al.* [127] purposely chose a virtual animal for their game that aims to teach children healthy behavior. The authors' choice was motivated by a survey of commercially successful mobile applications and the assumption that virtual animals do not exclude anyone by race or ethnicity.

Dormann *et al.* [63] state that the value of animal companionship in enhancing social competencies and psychological well-being is widely acknowledged. As shown by Chen *et al.* [46], virtual animal companions can support active self-reflection and learning in the affective and social domains. The importance of life-like animals is widely acknowledged in the therapy of phobias using VR applications, for example in the treatment of spider phobia as reported by Garcia-Palacios *et al.* [93]. Wrzesien *et al.* [341] generally discussed the potential impact of using VR and augmented reality (AR) technologies in therapies of phobia of small animals.

Previous work in HCI also explores the usage of virtual pets in particular. Ruckenstein [257], for example, suggests using virtual pets to encourage children to become more mobile in general. Altschuler [2] applies the theory of mind to recommend using virtual pets with autistic children to increase their appreciation for and consideration of others. Chen *et al.* [45] propose using virtual animals to encourage students to promote effort-making learning behaviors.

5.2.2 Animals and the Uncanny Valley

Previous work focuses on the relationship between familiarity (or eeriness) and the dimension of human-likeness; however, this does not explicitly state that the uncanny valley must have a dimension of human-likeness.

A study by MacDorman and Chattopadhyay [187] was conducted to determine whether *reducing realism consistency* in visual features increases the uncanny valley. Using transitions among features of realistic and CG entities, they showed that inconsistencies among features increased the eeriness for humans and animals (using bird and dog), however, not for objects. Corresponding to Mori's idea the authors hypothesized that "the more anthropomorphic the entity, the more reduced consistency in feature realism increases the uncanny valley effect" [187]. However, they found that inconsistencies between real and CG eyes, eyelashes, and mouth increased the perceived eeriness significantly *less* for humans than for animal entities, which could be different if more realistic animal models had been used.

Furthermore, Yamada *et al.* [342] showed that difficulties in categorizing human-like faces resulted in a negative overall impression. They attributed their results to the *categorization-based stranger avoidance* as well as the cause of eerie feelings in difficulties in categorizing an entity into a novel category, which theoretically could also apply to animals. They were criticized by MacDorman and Chattopadhyay [187] for artifacts in the used stimuli. In a position paper, Kawabe *et al.* [146] argued that the results are still consistent with the theory of categorization-based stranger avoidance. Furthermore, they conclude that the uncanny valley "can be extended to dimensions not directly related to human" even when Mori described it in a non-human to human continuum [146, 213]. A study by Bartneck *et al.* [15]. used a robot toy (iCat) to investigate how users perceive synthetic emotional facial expressions Their study indicates that emotional expressions of anthropomorphic entities are perceived categorically. In a response to this paper, MacDorman and Chattopadhyay [186] stated that this explanation is doubtful, for example, because "eeriness is seldom felt when meeting strangers."

Ferrey *et al.* [82] found trends depicting the uncanny valley that occurred for all continua including non-human stimuli and are hence not only related to humans. However, they used hybrid morph continua including human-like as well as animal-like faces to trigger unpleasant effects. Kawabe *et al.* [146] added that the eeriness occurring between different animal categories as used in the study by Ferrey *et al.* [82], is out of scope by the *realism inconsistency hypothesis* but could be well explained by the categorization-based stranger avoidance theory [186]. It remains unclear, whether this kind of eeriness as used by Ferrey *et al.* [82] can be attributed to the human-like or to the animal-like representation.

Ramey [238] concludes that the repulsion of uncanny depictions increases if the potentially applicable category of an observed character falls into the same category as the observer . This may lead to the assumption that there is a weaker uncanny effect on the perception of animal-likeness compared with human-likeness. Empirical evidence on explanations for the uncanny valley was summarized by Kätsyri *et al.* [145] in a meta-review. They found support in previous work for the *perceptual mismatch* hypothesis, which predicts that humans are more sensitive and less tolerant to deviations from typical norms when judging human faces [145]. However, they argue that the uncanny valley "is manifested only under specific conditions" and "that inconsistent levels of realism and atypical features represent different conditions leading to the uncanny valley" [145].

Reactions produced by artificial representations of realistic animals have been insufficiently considered so far. However, dead animals provoke similar emotions as the uncanny valley: disgust, revulsion, and fear [56]. Reasons for the uncanny valley have been associated with an innate fear of death and a subconscious strategy of coping with its inevitability [184]. There are also similar descriptions of ambivalent and uncomfortable feelings associated with dead animal bodies (*cf.* D'Zurilla [56] and MacDorman [184]). *Animal reminder* is one of three kinds of disgust described by Olatunji *et al.* [222]. MacDorman and Entezari [188] found Animal Reminder sensitivity is "confusing" but increases eerie ratings. Steckenfinger *et al.* [297] conducted an eye-tracking study with monkeys using CG renderings of monkeys and assume an evolutionary mechanism behind the uncanny valley.

5.2.3 Measurements and Operationalization

One measurement instrument to detect the subjective perception of humans towards robots is the Godspeed questionnaire proposed by Bartneck *et al.* [17]. The questionnaire is primarily designed for surveys in human-robot interaction (HRI). Special requirements in terms of the uncanny valley led to an improved alternative developed by Ho and MacDorman [121]. The questionnaire measures the subjective perception of people and has been successfully validated for robots and for virtual humans [190, 209]. To control the scale of human-likeness, Cheetham and Janke introduce linear blendings (morph continua) of CGI and photos to represent the dimension of human-likeness (DHL) [41]. Such blendings should reduce inconsistent findings when collecting subjective ratings of human-like stimuli. However, Kätsyri *et al.* [145] pointed out that the use of morph continua (*e.g.* as used by [41, 43, 185, 187]) between two images may produce blending artifacts. In this case, it cannot be precluded that the artifacts, not the categorical mismatch of the stimuli, cause negative responses.

5.2.4 Previous Findings Related to Animals

Previous work highlights the importance of virtual animals in games and in HCI [45, 127, 206]. Research of the uncanny valley investigates the effect of human-likeness using robots and virtual characters [183, 192, 238, 318]. Studies have used animal stimuli as subject of their investigations [82, 187, 297]; however, little is known about the potential existence of an uncanny valley for virtual animals. Measuring the effect is currently only established using questionnaires that are explicitly related to human-likeness [17, 190, 209], so an increased effect using motion as predicted by Mori could not be confirmed by previous work [213, 230, 308]. Previous research also used human-animal hybrids to investigate the uncanny valley [82]. In line with Ferrey *et al.* [82] and Kawabe *et al.* [146], we aim to deepen the understanding of eeriness in virtual animal perception.

5.3 Method Overview

Related work indicates that inconsistent levels of realism as well as atypical features separately cause the uncanny valley phenomenon. In particular, Kätsyri *et al.* [145] pointed out that there are possibly two different "Valleys". Therefore, we collected data about the subjective perception of humans towards virtual animals at different levels of realism as well as by adding unnatural features. We decided to use CG renderings (instead of linear blendings) of animals because they are well suited for controlling different degrees of realism.

Revealing the effect for one kind of animal is sufficient to determine whether the uncanny valley is applicable to at least one non-human species. This does not necessarily mean that there is the same effect for every kind of non-human virtual character or that the uncanny valley is generally applicable to virtual animals at all. Because the uncanny valley is related to familiarity, we assume that a potential effect of the uncanny valley on virtual animals rather occurs when an animal is familiar. Pets are familiar to humans, and cats are the most kept pets in many countries (see pet statistics from the U.S., U.K., and Germany[1]). Cats are used in video games and in HCI related research [344]. Inspired by the importance of cats in humans' everyday life, we assume that virtual cats are significant for many application domains, such as video games and health therapies. Thus, the following studies focus only on cats. We do not consider dogs, for example, due to the large physiological differences between breeds, which potentially lead to perceptual biases in humans. Effects using other animals are not investigated in the study.

We designed two studies using cat renderings as stimuli: In a first study, we collected quantitative measurements of the perceived eeriness towards renderings of virtual cats from high to low levels of realism. Supplementary to this, we collected qualitative feedback to gain deeper insights about the attitude of our participants towards virtual cats in current video games. As the related work shows, very realistic human-like faces with artificially enlarged eyes, facial expressions, and exaggeration of emotions [192, 194, 287, 318] cause very negative reactions. Thus, in a second study, we investigated these aspects using effects of atypically enlarged eyes and emotions on virtual cat renderings. To collect reliable data from a large sample, both studies were conducted using online surveys.

5.4 Study I: Realism and Eeriness of a Virtual Cat

In the first study, we aim to gain insights into humans' perception of virtual animals. Using a familiar pet species, namely cats, we aim to collect quantitative as well as qualitative feedback from a number of participants.

5.4.1 Study Design

The first study is divided into a quantitative and a qualitative survey. The quantitative part of the survey follows a repeated measures (RM) design with realism level as the only independent variable. We used the perceived realism, eeriness,

[1] Pet statistics from the year 2014: `http://www.petsecure.com.au` and `http://www.statista.com`

and aesthetic quality as dependent variables. In the qualitative part of the survey, we asked participants to provide feedback about their impression of four virtual animals from current video games and real-time applications.

5.4.2 Measures

Previous work investigated the effect of the uncanny valley using questionnaires that are explicitly related to human-likeness. Therefore, we developed a new questionnaire based on items from the successor of the Godspeed questionnaire introduced by Ho and MacDorman [17, 121]. Because perceived *human.likeness* and *attractiveness* are highly related to anthropomorphic (human-like) properties, we replaced these items by *naturalness* and *aesthetics*, respectively. Two semantic differentials are retained to evaluate the perceived similarity towards a real animal: *artificial-natural* and *synthetic–Real*. Two further semantic differentials ask for *familiarity*: *uncanny-familiar* and *freaky-numbing*. The last two semantic differentials are adopted to evaluate *aesthetic* aspects: *ugly-beautiful* and *unaesthetic-aesthetic*. All items were rated on a 7-point Likert-scale. We operationalize familiarity ratings as the measure to test the hypotheses.

5.4.3 Stimuli

We opted to use renderings of a short-haired cat in a neutral pose. To avoid prejudices caused by white or black furred cats, we chose a gray-haired cat (Russian Blue) without using any prominent fur pattern. However, fur does not allow linear morph continua as used by Cheetham *et al.* and Ferrey *et al.* [41, 82]. Therefore, we developed seven virtual cat models with a gradually increased level of realism. All images (R1-8, depicted in Figure 5.3) are based on a reference photo (R1).

The reference photo shows a gray short-hair Russian Blue cat sitting in a neutral pose without disturbing artifacts or distracting background. We use this image because the scene setup is neutrally arranged and the whole scene is easy to reconstruct in 3D. Textures and models were created in Autodesk Mudbox by a professional CG artist. Cat, background, area lights, and camera were reconstructed in Autodesk Maya. All 3D images (R2-7) were rendered using global illumination and ambient occlusion. Fur of the high-quality (HQ) model

Figure 5.2: CG rendering (Stimuli R2) of a virtual cat with realistic fur.

(R2, see Figures 5.2 and 5.3b) was created using hair geometry with the XGen-Tools for Maya. The cat model in image R2 has 134,256 triangles. The following image (R3) was rendered using a simplified and noise-free hair model. Hair geometry was removed and polygons were reduced for the next two images (R4, 16,782 triangles and R5, 8,390 triangles) corresponding to levels of realism used in current video games. The model as well as the texture quality of the ultra low-poly model (R6, 424 triangles) were further reduced to an ultimate minimum as used in games for very low levels of details. The model of the second to last image

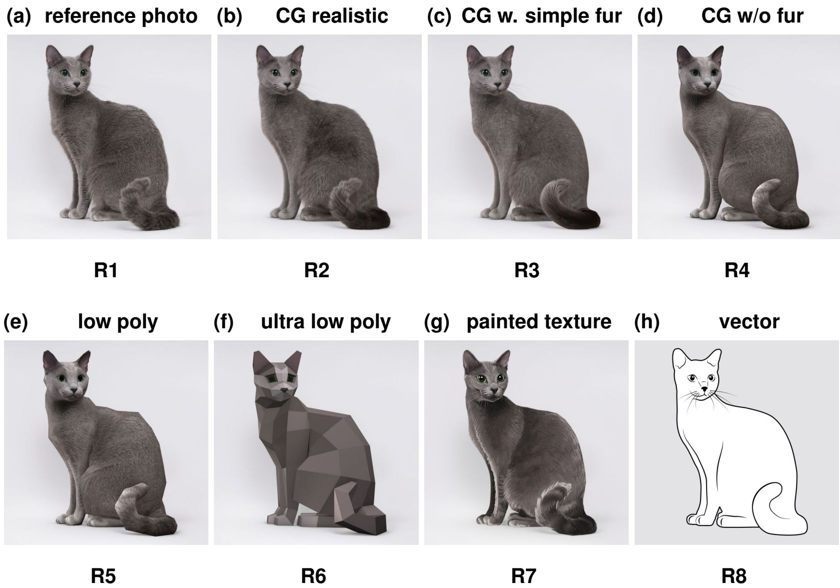

Figure 5.3: Cat depictions in different levels of realism. The first image R1 (a) is the reference photo (by kuba_girl / Shutterstock.com). R2-7 (b-g) are rendered in 3D. R2 (b) contains complex fur geometry, R3 (c) contains simplified fur. R4 (d) and R5 (e) are textured cat models as they are used in video games. R5 (e) contains no subdivisions. R6 (f) shows a more simplified facet model without textures. R7 (g) was rendered in toon style. R8 (h) is a contour enhanced vectorization without colorization.

(R7) received a hand-painted texture and toon-look using Cel-shading. The render size corresponds to the original resolution of the reference image (2456×2718 pixels). The last image (R8) was manually vectorized in Adobe Illustrator 6 with simplified contours in black and white. All images show the same cat in the same pose using different levels of realism. Using these stimuli, we assume that we are able to transfer the paradigms from CG humans to CG animals.

For the second part of the study, we took in-game screenshots of the three video games *Grand Theft Auto V*, *The Witcher III - Heart of Stone*, *The Sims III* as stimuli. The images are shown in Figure 5.4. These video games were selected based on their topicality and dissemination. An additional image is a rendering

Figure 5.4: Virtual cats from recent video games and real-time application: (a) *Grand Theft Auto V*, (b) Nibbles from *The Witcher III - Heart of Stone*, (c) standard model of a Russian Blue Cat from *The Sims III + Pets Expansion Pack*, (d) commercially available low-poly cat from `turbosquid.com`.

of a commercially available real-time 3D model, which was downloaded from an online store for 3D Models (turbosquid.com). All screenshots were captured using lossless quality settings and were cropped to an image size of 500×400 pixels. No image was scaled.

5.4.4 Survey Procedure

Participants obtained a link to our survey website. They were asked to maximize their browser and received information about the procedure. After collecting demographic data about gender, age, pet ownership, and home country, the first stimulus was presented in full-screen within the browser. Clicking on an image reduced its size to 30% and revealed the rating scales. A participant could only proceed when all questions were completed. By clicking again, the image size could be increased back to full-screen. The rating scales consisted of the six word pairs based on the bipolar scales in terms of realism, familiarity, and aesthetics. The scales were randomly sorted and oriented to avoid biases. The eight stimuli were presented in random order.

After finishing the first part of the study, the participants were asked to provide comments about the virtual cats of current real-time applications (see Figure 5.4), and to answer the following questions in separate input fields: *Describe in your own words your personal impression about the depiction of the cat above. Please give reasons for your impression. Are there features that particularly attract your attention?* The participants could continue without leaving a comment. Finally, the participants were invited to leave their e-mail address to take part in a draw for a gift card. The mean survey completion time was 12.5 minutes ($SD = 10.1$).

5.4.5 Participants

We recruited participants via Facebook, Twitter, online forums, and mailing lists of two universities in Germany. In total, 339 participants (152 males, 186 females, 1 other/not specified) from age 15 to 76 ($M = 24.7$, $SD = 6.6$) completed the study. The sample includes 44.8% males, 54.9% females, and .3% other/not specified. Home countries of the participants corresponded to the demographics of our university's undergraduate population (93.8% from the German speaking area, 6.2% foreign or unclassified). Ninety-three participants (31.78%) stated that they currently own at least one cat as a pet. Seventy-six additional participants (22.41%) pointed out that they owned at least one cat as a pet in the past. This means that 169; almost half of all the participants (49.8%), had currently or previously owned a cat.

5.4.6 Quantitative Results

To assess the reliability of the three measures we conducted a Spearman-Brown correlation analysis, which is considered as the most reliable estimate for two-item scales [71]. The correlation matrix shows that items within realism ($\rho = .836$, $p < .001$), familiarity ($\rho = .715$, $p < .001$), and aesthetics ($\rho = .830$, $p < .001$) have the highest correlations among the measures (all others with: $\rho \leq .709$, $p < .001$).

We conducted a repeated measure (RM) one-way multivariate analysis of variance (MANOVA). Using Wilks' lambda, there was a significant effect of stimuli realism on our three measures, $\Lambda = .062$, $F(21,318) = 227.703$, $p < .001, \eta^2 = .399$. Separate univariate analyses of variance (ANOVAs) on the dependent variables revealed significant effects on familiarity, $F(7,2366) = 284.738$, $p < .001$), perceived realism, $F(7,2366) = 722.195$, $p < .001$), and aesthetics, $F(7,2366) = 226.235$, $p < .001$. Figure 5.5 shows the results of the three measures collected in the first part of the study.

Bonferroni corrected pair-wise t-tests revealed significant differences between all conditions ($p < .05$) except between R3–R8 ($p = 1$), and R5–R6 ($p = 1$) for ratings of familiarity and between R1–R2 ($p = .375$), R4–R6 ($p = .056$), R4–R7 ($p = .309$), and R6–R7 ($p = 1$) for ratings of aesthetics. All pair-wise comparisons of perceived realism were significant (with all $p \leq .002$). From the stimuli R1 ($M = 5.674$, $SD = 1.176$) to R5 ($M = 2.948$, $SD = 1.265$), we found a significant decline of the subjectively perceived familiarity. From R6 ($M = 2.988$, $SD = 1.282$) to R8 ($M = 4.798$, $SD = 1.270$) the results indicated an ascent of familiarity again. The mean ratings of R5, R6, and R7 are below the neutral average of 4.0. The subjectively perceived aesthetics decrease from R1 ($M = 5.856$, $SD = 1.095$) to R5 ($M = 2.827$, $SD = 1.285$) and increase again from R6 ($M = 3.870$, $SD = 1.663$) to R8 ($M = 4.556$, $SD = 1.330$). As mean ratings of familiarity, mean ratings of aesthetics are below the neutral average of 4.0 for R5, R6, and R7. Measurements of the subjectively perceived realism decrease from R1 ($M = 6.087$, $SD = 1.143$) to R6 ($M = 1.933$, $SD = 1.030$) and increase with R7 ($M = 3.706$, $SD = 1.152$) and R8 ($M = 2.591$, $SD = 1.448$). All data analysis was performed using SPSS v21.

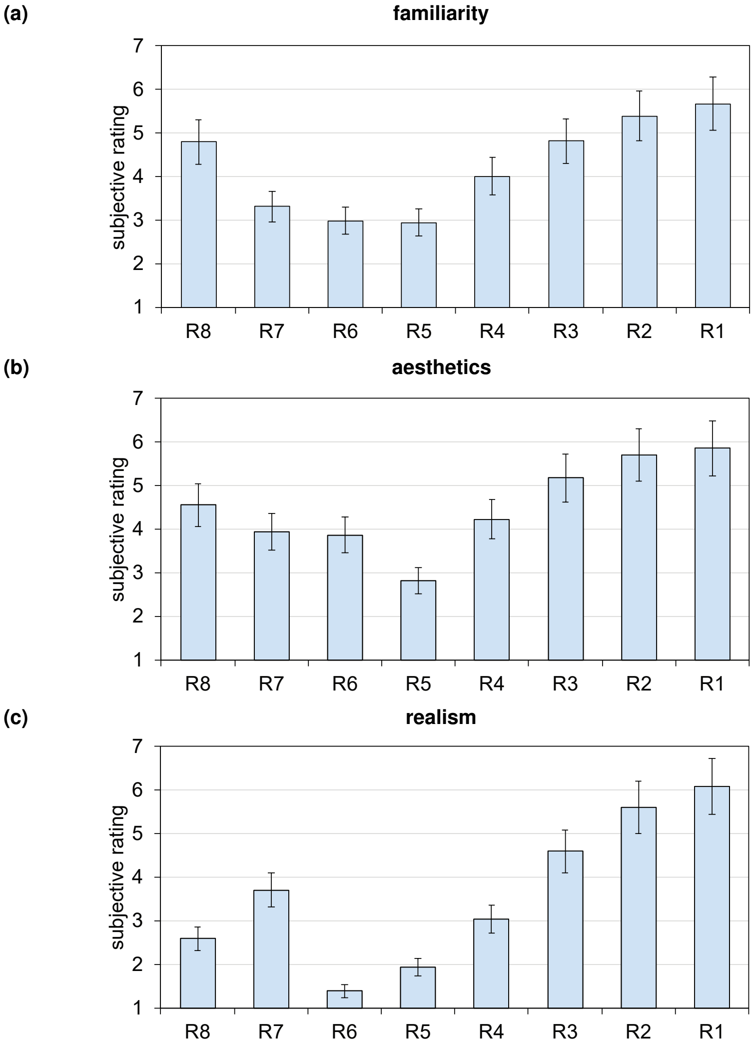

Figure 5.5: Mean subjective ratings of perceived (a) familiarity, (b) aesthetics, and (c) realism. Ratings of the eight stimuli sorted in the realism continuum ranging from R1 (real) to R8 (abstract). Error bars show the 95% CI.

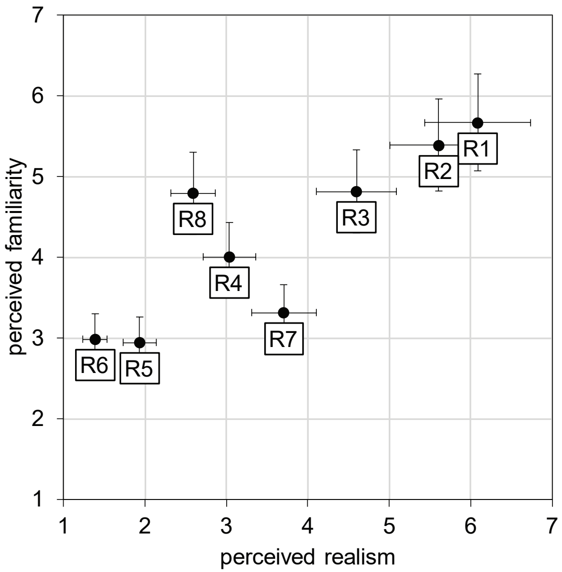

Figure 5.6: Subjective perception of realism related to familiarity of the stimuli R1 (real) to R8 (abstract). Error bars show the 95% CI.

The interpretation of the results depends on the order of the eight stimuli. The stimuli represent a constructed continuum from most realistic (R1) to least realistic (R8) CG renderings. Using this order, the results seem to shape a "valley" of familiarity. However, Figure 5.6 shows the familiarity ratings related to the assessments of realism. Through reordering the stimuli based on the respondents' assessments of realism the position of R7 moves to between R3 and R4, and the position of R8 moves to between R4 and R5.

As previously mentioned, almost half of all participants ($N = 169, 49.8\%$) have had a cat in their life. Participants who stated that they currently or ever had at least one cat as a pet in the past were summarized as cat owners. Moreover, there was a balanced number of male and female participants. Potential effects of both factors were investigated by conducting a RM-MANOVA including cat ownership and gender as between-subjects variable. We found significant effects of gender, $\Lambda = .973$, $F(3,333) = 3.048$, $p = .029$, and cat ownership, $\Lambda = .966$,

$F(3,333) = 3.894$, $p = .009$. There was no interaction effect between both factors, $\Lambda = .979$, $F(3,333) = 2.400$, $p = .068$, however, a significant interaction effect of gender × realism $\Lambda = .899$, $F(21,315) = 1.687$, $p = .031$.

Through separate univariate ANOVAs, we found significant effects of gender on perceived familiarity, $F(1,335) = 5.064$, $p = .025$, and significant effects of cat ownership on familiarity, $F(1,335) = 6.061$, $p = .014$ as well as on aesthetics, $F(1,335) = 4.873$, $p = .028$, (all others with $p \geq .566$). We found significant interaction effects of gender × realism on eeriness, $F(7,2345) = 3.373$, $p = .001$, and aesthetics, $F(7,2345) = 2.842$, $p = .006$, but not on perceived realism, $F(7,2345) = 1.759$, $p = .247$. The tests revealed that familiarity was rated significantly higher by men ($M = 4.330$, $SE = .051$) than by women ($M = 4.174$, $SE = .046$); however, this depended on the degree of realism. Female participants showed lower familiarity ratings at lower degrees of realism. Cat owners' perceived familiarity was significantly higher ($M = 4.337$, $SE = .049$) than non-cat owners ($M = 4.167$, $SE = .049$). Aesthetics ratings were significantly higher for cat owners ($M = 4.619$, $SE = .052$) than for participants who never had a cat ($M = 4.431$, $SE = .011$).

5.4.7 Qualitative Results

Participants provided 786 comments in total: 199 about the cat from *Grand Theft Auto V* (in the following labeled as GTA5), 195 about the cat from *The Witcher III: Heart of Stone* (WITCHER3), 186 about the cat from *The Sims III + Pets Expansion Pack* (SIMS3), and 206 about the commercially available model from *turbosquid* (TS). Participants were asked to describe their personal impression of the depiction. They were also asked which features especially attracted their attention. All comments were analyzed and coded. Two researchers went through all transcribed notes to check each other's coding and to establish consistency in the assignment of codes to the same phenomena. With open coding, the first iteration of grounded theory, three main categories of perception in virtual animals were found: animal-likeness, mostly scene related aesthetic qualities, and facial as well as body expressions.

Naturalness: The first category is a mental comparison with the *naturalness* of a real animal. Previous knowledge of what an animal has to look like may lead to a perceptual mismatch when regarding an animal that does not look completely natural. Ambiguity, missing, or incorrect attributes may lead to a negative impression of a virtual animal depiction. To describe the lack of naturalness, participants often used other animals (or even objects) to describe their impression. *“Seems like a stiff plastic toy. The fur does not look fluffy and the eye glance is too strong”* (SIMS3); *“The cat does not look like a cat – more like a hyena, because of the short, curved back”* (TS); *“The cat looks more like a panther [...]. In general, the colors are too intense to be considered realistic”* (SIMS3); *“The cat has legs like a dog [...] or the body of a wild cat”. “Looks rather like a rat”* (GTA5). *“Looks like a robot. Artificially, cannot really rate it”* (SIMS3). *“The cat looks a bit like a dog"* (SIMS3); *“The cat looks like a mummy”* (TS); *“Scary, because the cat looks more like a mutant”* (TS); *“The object is more similar to a raccoon than to a cat.* (GTA5). Also missing features that make the comparison with a natural cat difficult were considered negative by the participants: *“Missing fur makes the ’fluffiness’ and thus the very likeable aspect of cats disappear”* (SIMS3); *“No whiskers. Ears are a bit too long. Hair is missing. Looks more like the physique of a dog. Face of the cat is ugly”* (TS).

Aesthetic qualities and relation to the scene: A negative impression might be reinforced by the lack of aesthetic qualities which were often brought into connection with the scene. Therefore, we summarized aesthetic qualities and how the cat fit into the scene in a single category. These aspects are also influenced by properties of the environment and are related to how the animal fits into the scene. One example is the lack of shadows, which gives the impression the cat is levitating: *“Cat levitates in the air. [...] Dead eyes and ears look like they are clipped out. Although quite realistic, but also quite unaesthetic”* (TS). *“Good lighting, but missing details. Shadow isn’t correct. Cat seems to levitate”* (WITCHER3). *“Posture and missing fuzz from the fur make it look less realistic”* (WITCHER3). *“The cat could be more realistic. Looks like it is pasted into the image. The size of the cat in relation to the environment bothers me”* (GTA5). *“The cat looks like it is cut out and glued on”* (GTA5). *“The dark shadow makes the* *cat look creepy. It also acts vigorously and pugnacious. Certainly in this case the environment plays an important role"* (WITCHER3); Furthermore, the overall look of a cat depiction was considered negatively: *"The cat's appearance is stylized rather than realistic. Textures and colors seem slightly exaggerated"* (SIMS3). *"The patterning makes it look real. The eyes are not bad. The uncanny part comes as you combine this realistic looking texturing with a low res model"* (SIMS3).

Health status and body language: The participants could not comment with any certainty about the cats' health status or their body language. This also includes facial expressions, angry eyes, and aggressive body postures. We found that the participants inspected the cat depiction carefully to see if it could be a threat or a disease carrier. We summarized these comments into one category because of their relation to an evolutionarily related explanation (contact avoidance due to uncertainties). Considering rabies where the infected animal is aggressive and attacks without provocation, both health status and body language pose a threat: *"The cat looks as if it has a bad disease"* (TS); *"The cat looks scary, because I cannot understand what its body and facial expressions really mean"* (TS); *"Cat in aggressive posture, nasty facial expression"* (GTA5); *"Cat has an aggressive attitude, the facial expression looks uncomfortable"* (WITCHER3); *"Looks like a statue, because the attitude is very symmetrical and unnatural"* (SIMS3); *"The facial expression is too rigid, the body very voluminous. Therefore, the cat looks a little bit scary to me"* (GTA5); *"The fur looks very good, good posture, attitude, and snout. Only the eyes are scary"* (WITCHER3); *"Facial expression looks artificial because of the forehead"* (WITCHER3). In particular, the appearance of the eyes was mentioned and emphasized in contrast to other body parts: *"Evil eyes. Belligerent"* (GTA5); *"The eyes of the cat look scary. Otherwise, the body is well done"* (WITCHER3).

Summary

When a virtual depiction contradicts the familiar concept of an animal, a negative impression arises. We consequently derive the following triggers for the violation of a familiar virtual animal depiction: Mistakes in natural appearance, unaesthetic aspects of the animal within a scene, and a threatening or rigid body as well as facial expressions. Only when these attributes are considered positive, willingness for interaction arises. *"No realistic proportions and the face is uncanny. Thus, this cat is not cuddly"* (TS). Furthermore, people feel threatened by concerns about the health status or body language. Avoiding direct contact with an animal whose health condition is suspect or which looks unfriendly may have an evolutionary purpose.

Some comments point to certain expectations or habituation towards computer game graphics. *"The cat reminds me of a computer game"* (TS). Interestingly, the virtual cats are dated much older than the game from which they originate. The oldest video game from our selection is from 2011 (Sims III) *"Reminds me of old games like Tomb Raider or Sims I"* (SIMS3). *"Looks like 10-year-old computer game graphics"* (WITCHER3). *"A bad computer cat out of the 90s"* (SIMS3). *"I get nostalgic about old computer games"* (SIMS3). *"Asset from the 80s?"* (TS). This means that virtual animals may trigger a perceptual shift to older game graphics even though the game is more recent.

5.4.8 Discussion

In the first study, we collected quantitative ratings as well as qualitative feedback. In the first part of the study, we used a reference photo and seven CG images with a varying degree of realism to measure the perceived familiarity of a virtual cat. High levels of realism were assessed to be very familiar. Lowest ratings for familiarity were measured for realism levels as used in current video games (see Figures 5.3e and 5.3f). Stylized and unrealistic levels of a virtual cat received higher ratings of familiarity (see Figures 5.3g and 5.3h) again. Thus, we found a decrease of familiarity using cat depictions at intermediate graphic levels of realism, which results in a long U-shaped valley (see Figure 5.5). This is predicted by Mori's hypothesis of the uncanny valley and was verified for virtual human-like characters [192, 213, 318].

The shape of the valley would differ if the stimuli were sorted by realism ratings of the participants. Figure 5.6 shows R1–R6 in an almost linear relation among familiarity and realism, while stimuli R7 and R8 are not part of this linearity. The manipulation of realism could be compromised due to the following reasons: (1) the concept of realism is partially biased by other associations; and (2) stylized or abstract images might not belong to the same continuum of realism. The kind of stimuli changed from realistic (R1–R6) to another perceptual construct (R7, R8) which is finally considered as non-real anymore. Thus, the results cannot clearly confirm H1. We assume that the same problem of using one single continuum of realism exists for both human-like as well as animal-like characters. Reducing the degree of realism by adding abstraction does not consequentially mean to map points on the same continuum. However, this does not explicitly contradict Mori's hypothesis, that brought multiple categories ("industrial robot", "stuffed animal", "zombie") into a related continuum ("human-likeness"). We will later discuss how difficulties of categorical perception can be integrated into the theory of the uncanny valley of animals. Further problems with the dimension of human-likeness and the usage of gradual continua (such as artifacts) are discussed by Kätsyri *et al.* [145].

Furthermore, the results show that the virtual cat at intermediate (R4, R5) and not at high photo-realistic levels of realism (R5–R7) is rated as less familiar than the photo (R1), toon painting (R7), or the simplified vector illustration (R8). Instead of a sudden decrease of familiarity, the familiarity ratings rather show a downhill slope between R1 and R6. This could be potentially caused by a shift of familiarity due to a higher sensitivity towards the own (and more familiar species). This was not supported by our results, as perceived familiarity due to cat ownership was not significantly affected. Furthermore, using a very large set of robot faces Mathur and Reichling [195] found that the deepest point of the valley can potentially found at very intermediate and not at very high levels of human-likeness as Mori predicted [195, 213]. Therefore, we suggest that future work should directly compare human and animal entities to investigate this shift.

Results of the first part of the survey show significant differences of the perceived realism between all stimuli. This means that the reference photo (R1) and the CG rendering with the highest level of virtual realism (R2) can still be distinguished from each other. However, ratings of familiarity and aesthetics of a high-level advanced CGI model (R2) do not significantly differ from the reference photo. Therefore, we assume that virtual animals can be rendered at realism levels where they receive high acceptance. This is confirmed by the current trend of movies using very realistic computer-animated animals. However, advanced rendering and post-processing techniques, as used in animated movies, are currently not applicable to animals in video games, which indicate that they are currently affected by the uncanny valley. The uncanny valley hypothesis, however, predicts that *almost* realistic characters fall into the valley. Therefore, it needs to be discussed, *where* the uncanny valley is in animals. In particular, R6 has low realism scores; however, is clearly not close to the real cat. This suggests that R6 is potentially on the left or "safe" side in Mori's graph. Higher ratings of familiarity and aesthetics of cat owners and no interaction effects reveal that having at least one cat as a pet generally improves the participants' attitude towards the cat depictions. However, we found no enhancement of the amplitudes (lower ratings at lower degrees of realism; higher ratings and higher degrees of realism), which could indicate that familiarity increases the perceptual sensitivity towards an animal entity.

Qualitative results show that the graphical standard of current computer games may lead to an uncanny perception of animals. Comments regarding current animals in real-time environments indicate, in addition to an individuals' attitude towards animals, three different factors to be responsible for an eerie sensation: The naturalness of an animal model, its expression, and its aesthetic qualities which are strongly influenced by the scene environment (*e.g.* lights). We found that people seemed to be confused and rated virtual cats negatively when they perceived the depiction as ambiguous. Negative responses due to violations of the expectation regarding an animal's outward appearance would support the perceptual mismatch hypothesis [41, 342].

We therefore assume, in respect to games, that when a player pays attention to other aspects of the game (scene, story, etc.), the visual quality of the rest of the scene may cover the potentially eerie appearance of an animal. However, if a virtual animal with a slightly abnormal appearance or unusual expression of face and body is the focus of a game's scene, its depiction may leave a negative impression. Qualitative feedback also reveals that participants were partially reminded of older video games and older graphical standards. We assume that there are effects of habituation or expectation that may have an influence on the perception of virtual depictions. These aspects could also have an influence on the uncanny valley.

5.5 Study II: Effects of Stylization and Emotions

The qualitative part of the first study indicates that unusual expressions and an exaggerated appearance may increase the eerie effects of a virtual animal. These aspects were not considered in the first study. The related work shows that very realistic human faces with atypical features such as artificially enlarged eyes cause very negative reactions [192, 287]. Two further studies show that facial expressions and exaggeration of emotions cause larger effects if the face is more human-like [194, 318]. We used anthropomorphic emotions due to their frequent usage in current animated movies as in *The Smurfs I* (2011) [100] or *The Jungle Book* (2016) [76]. Anthropomorphic emotions as well as artificially enlarged eyes are atypical features of cats, which are further investigated using different levels of rendering in the second study. Thus, we aim to investigate and determine whether and how atypical properties can be transferred to virtual cat renderings.

5.5.1 Study Design

We reduced the number of realism levels to four to reduce the overall number of possible conditions selecting R1, R2, R4, and R5 from the first study. The realism level of R3 was excluded due to only minor structural changes of the cats' fur compared to R2. Abstraction or abstract-like levels of realism as used in R6, R7, and R8 were excluded as well, so as to model a single continuum of the cat's realism (R1–R4). In addition to the general neutral style of the cat, we added atypical features through enlarging the eye size and giving the cat a stylized appearance. Facial changes expressed three states of emotion: neutral, happy, and sad. Thus, for the second study, we used 4 levels of REALISM, 2 levels of STYLIZATION, and 3 levels of EMOTION in a multi-factorial within-subject design. We used the same measures of familiarity, realism, and aesthetics as introduced in the first study.

5.5.2 Stimuli

A matrix of 24 stimuli based on the cat depiction from the first study (see Figure 5.3) was created. Four levels of realism were combined with 2 levels of stylization and 3 levels of emotions resulting in these 24 conditions (close-ups of all stimuli see Figure 5.7). The cat was given a stylized appearance through enlarging the eye size (140%) and two contrasting emotions (happy and sad). The 3D models used in the first study were morphed and rendered again. The photo reference was manipulated using Adobe Photoshop.

5.5.3 Survey Procedure

Participants obtained a link to our survey, where they received information about the survey and the terms of use. After collecting demographic data about gender, age, and game as well as video usage, the stimuli were presented. As in the first study, participants rated each image using six word pairs on a bipolar seven-point scale in terms of realism, familiarity, and aesthetics. All corresponding adjectives were randomly sorted to avoid biases. Orientation, as well as the order of the scales, were randomized and placed after each image. The order of images was randomized as well. An image change, which was initiated by clicking on the next-button, appeared with a delay of two seconds to avoid direct comparisons. The average time to complete the survey was 33.1 minutes ($SD = 28.33$).

5.5.4 Participants

Participants of the second online survey were also recruited via Facebook, forums, and mailing lists of our two universities in Germany. In total, 214 participants, (91 male, 121 female, 2 other/not specified) took part in the second study. The sample include 42.5% male, 56.5% female, and .9% other/not specified participants. Participants age ranged from 18 to 44 ($M = 23.29$, $SD = 3.77$). Home countries of the participants reflected the demographics of our university's undergraduate population (92.5% from the German speaking area, 7.5% foreign or unclassified). 68 participants (31.78%) stated that they currently had at least one cat as a pet,

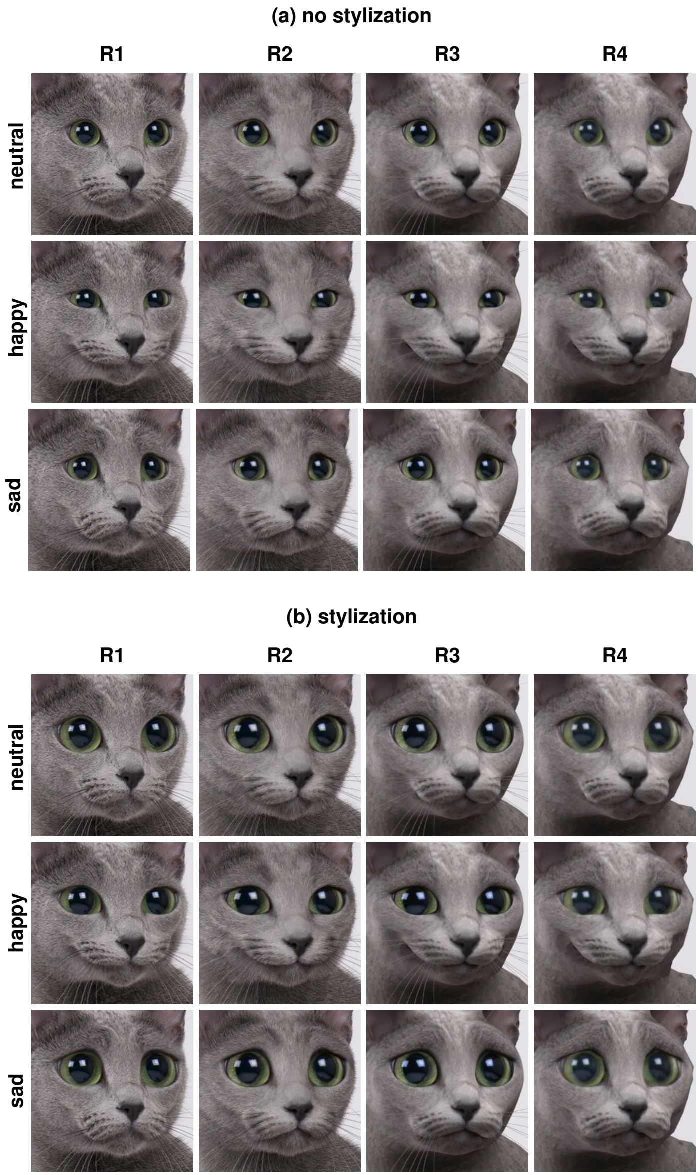

Figure 5.7: Close-ups of the virtual animals renderings. We used three factors in the second study, 4 levels of realism (R1–R4) were combined with 3 levels of emotions and 2 levels of stylization which results in 24 different conditions.

49 additional (22.90%) participants that they had owned at least one cat as a pet in the past. This means that 117 participants (54.67%) had owned a cat as a pet in their lifetime.

5.5.5 Results

As in the first study, the reliability of the three measures was assessed using Spearman-Brown correlation analysis. The correlation matrix shows that items within realism ($\rho = .822, p < .001$), familiarity ($\rho = .728, p < .001$), and aesthetics ($\rho = .789, p < .001$) have the highest correlation among other items (all others with $\rho \leq .713, p < .001$).

A $4 \times 2 \times 3$ RM-MANOVA was significant for REALISM, $\Lambda = .156$, $F(9,205) = 123.052$, $p < .001$, $\eta^2 = .844$, STYLE, $\Lambda = .238$, $F(3,211) = 224.834$, $p < .001$, $\eta^2 = .762$, and EMOTION, $\Lambda = .283$, $F(6,208) = 87.658$, $p < .001$, $\eta^2 = .717$. There were significant interaction effects of REALISM $\times$ STYLE, $\Lambda = .296$, $F(9,205) = 54.165$, $p < .001$, $\eta^2 = .704$, REALISM $\times$ EMOTION, $\Lambda = .418$, $F(18,196) = 15.187$, $p < .001$, $\eta^2 = .582$, STYLE $\times$ EMOTION, $\Lambda = .420$, $F(6,208) = 47.861$, $p < .001$, $\eta^2 = .580$, and REALISM $\times$ STYLE $\times$ EMOTION, $\Lambda = .534$, $F(18,196) = 9.483$, $p < .001$, $\eta^2 = .466$. Separate univariate ANOVAs for each dependent variable revealed significant effects as well as significant interactions of the three factors. The results of the factorial analysis are listed in Table 5.1. Similar to the first study, Bonferroni-corrected pairwise comparisons between the levels of realism revealed significant differences ($p < .05$) except for the levels R1 and R2 for familiarity ($p = .984$) and aesthetics ($p = .666$). Post-hoc tests between all three emotional states revealed significant differences ($p < .001$).

We analyzed the familiarity rating to understand a participant's affinity towards stylization and emotional expressions on virtual animals. As previously mentioned, post-hoc comparisons revealed significant differences between the total means of all three emotional states (with all $p < .001$). Figure 5.8 shows that adding stylization using enlarged eyes as well as emotions strongly decreases the familiarity between all kinds of non-stylized depictions. The unchanged reference photo was rated with the highest familiarity. Facial expressions as well as stylization negatively influenced the perceived familiarity. We also found significant

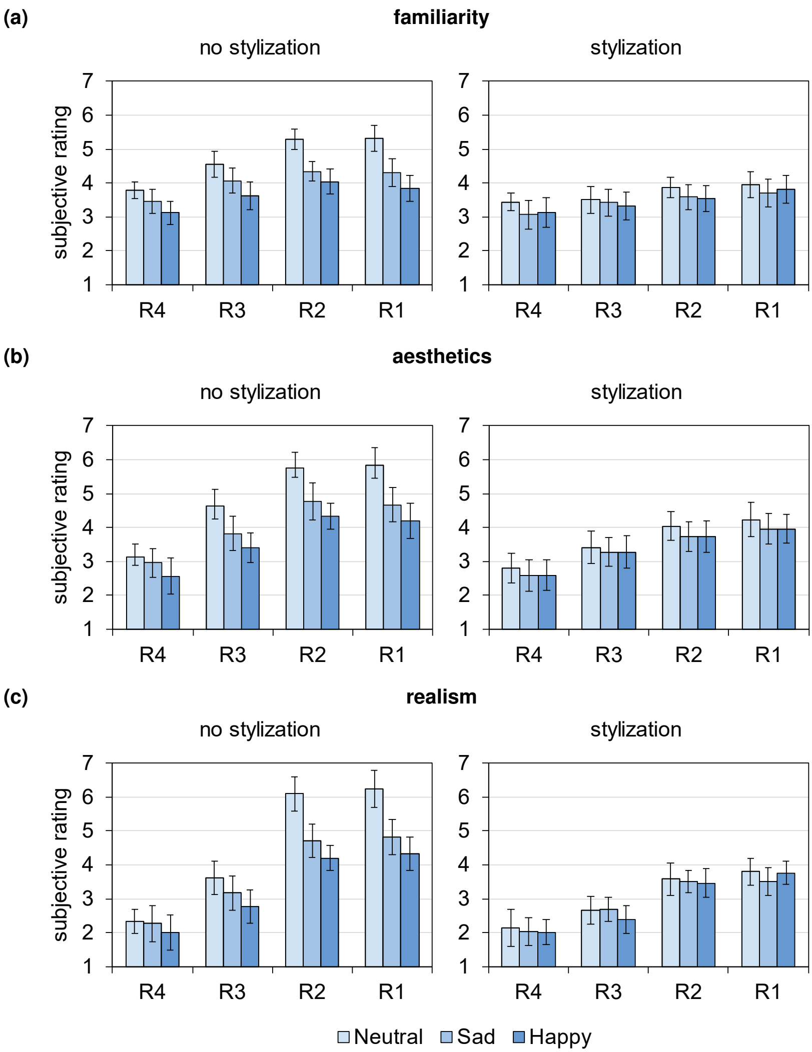

Figure 5.8: Subjective virtual animal ratings of perceived familiarity (a), aesthetics (b), and realism (c) for each emotional facial expression of the cat (neutral, sad, happy), separated by stylization. Error bars show standard deviation.

Factor	df	error	realism F	familiarity F	aesthetics F
REALISM	3	639	612.049 *	230.363 *	430.828 *
STYLIZATION	1	213	668.104 *	320.292 *	291.392 *
EMOTION	2	426	212.792 *	182.055 *	160.927 *
REALISM×STYLIZATION	3	639	163.471 *	52.062 *	53.806 *
REALISM×EMOTION	6	1278	27.232 *	13.022 *	12.785 *
STYLIZATION×EMOTION	2	426	135.614 *	113.325 *	80.656 *
REALISM×STYLIZATION×EMOTION	6	1278	24.090 *	8.581 *	5.651 *

Realism error df = 639; Stylization error df = 213; Emotion error df = 426, * for all: p < .001,

Table 5.1: Main and interaction effects of three RM-ANOVAs. The results show the effects of realism, stylization, and emotions on perceived realism, familiarity, and aesthetics.

differences between all states of emotional expressions. In most of the conditions, except for the stylized photo reference of the cat, the sad expression was rated more familiar than the happy facial expression.

As presented in Table 5.1, interaction effects between emotions and stylization as well as between emotions and realism were confirmed (with both $p < .001$). No significant effects of cat ownership ($p \geq .132$) and no significant effects of gender were found ($p \geq .076$). There were no further significant interaction effects ($p \geq .119$), except for style×cat ownership, $\Lambda = .899$, $F(3,101) = 3.778$, $p = .013$. No further effects of gender and cat ownership were found using univariate tests.

5.5.6 Discussion

The second study shows that two findings of uncanny valley research can be potentially transferred from humans to animals: We found significant interaction effects between all three factors: realism, atypical features given by stylization, and emotion. Decreasing realism or adding atypical features (enlarged eyes) lead to significantly decreased familiarity ratings of a virtual animal from stimuli R1 to R4. As the related work shows, this is also the case using humans or human-like characters (*cf.* Seyama *et al.* [287] and MacDorman *et al.* [192]). Through decreasing realism and adding anthropomorphic facial expressions on a virtual animal familiarity decreases as well. This aspect is confirmed by the related work using human-like characters, too (*cf.* Mäkäräinen *et al.* [194] and Tinwell *et al.* [318]). Finally, we found a statistically significant interaction effect between all three measures, which means that the factors are interrelated and influence each other when combined. We conclude that using less virtual animal realism atypical features decreases familiarity, which supports H2.

It is conceivable that the effect appears because of using human-like and not animal-like facial expressions. Smiling, for example, reduces the familiarity of an animal, because it is a typical human facial expression and does not reflect real animal behavior. This means that an animal character should not possess any human expressions, emotions, or speech if it should be fully accepted. At lower levels of realism, the difference between neutral depictions and depictions with emotions is not as large as at the difference between both at higher levels of realism. This means that virtual animals rendered at a high level of realism should not deviate from their natural appearance. In our first study, we found small effects of gender and cat ownership. These results were not confirmed in our second study. We compared the results with the first study and found lower familiarity and aesthetics ratings of female participants towards CG characters at lower degrees of realism (R5–R8), which were not used in the second study. However, it is conceivable that ratings between male and female participants potentially depend on the species.

5.6 Discussion

5.6.1 Summary

In this work, we conducted two studies to investigate whether the uncanny valley, originally developed for human-like characters, applies to virtual animals (RQ3). The focus of our investigation is to determine whether and when virtual animals cause negative reactions at certain levels of realism. Our studies concentrate specifically on depictions of cats as they are relatively familiar to humans and frequently used in video games and animated movies.

In our first study, we conducted an online survey to investigate how cats rendered at different levels of realism are perceived. We found a decrease of animal familiarity at intermediate but not at higher rendering realism. Compromised realism manipulation either indicates that the measured construct of realism is potentially affected by other constructs and could be invalid or that stylized and abstract stimuli respectively are part of another construct. Furthermore, the decrease is not as steep as predicted in Mori's graph for human-like characters. Qualitative feedback from participants judging depictions of cats in current video games supports an uncanny valley of animals and reveals potential causes. We identified three major factors that affect humans' reactions: the violations of the naturalness of the virtual animal, the facial expression and body pose, and how the animal fits into the scene.

Based on our results of the first survey and indications in previous work, we conducted a second study. We investigated the effect of stylization and facial expressions to substantiate further causes for the negative perception of a virtual animal. The results of the second study show that emotions and stylized appearance have a larger effect on familiarity at higher levels of realism compared to lower realism levels. Violations of the expected appearance of an animal cause negative reactions. We assume that our results are potentially caused by the hypothesized uncanny valley by Mori and that previous findings of research investigating this theory can be transferred from humans to depictions of at least one animal [213].

5.6.2 Integration into Existing Theoretical Frameworks

In line with Ferrey *et al.* [82] and Kawabe *et al.* [146], we suggest to examine an extension of Mori's hypothesis to a broader definition of the uncanny valley by including non-human entities. Nevertheless, how does this new aspect fit into Mori's theory and how can this be explained? The extension of the theory to animals is theoretically applicable to different recent explanations of the uncanny valley: Categorical perception and perceptual mismatch are not explicitly restricted to human-likeness.

Difficulties in categorical perception evoke negative responses [32, 33, 40]. It is conceivable that our results are caused by negative ratings due to difficulties while discriminating between animals and non-animals. In our first study, we found an almost linear relationship for familiarity and realism between the real and CG animal stimuli (R1–R6). A high sensitivity towards violations of a known concept (cat) potentially leads to a categorical discrimination between a natural and an abstract entity. Visual cues in the abstract category are assessed by individual and aesthetic preferences of the participants. Difficulties in discriminating abstract and real animals might lead to similar negative familiarity ratings as observed for human-like characters [32, 40, 43, 82]. As previously mentioned, Mori brought multiple categorical entities ("industrial robot", "bunraku puppet") into one continuum of "human-likeness."

In scope of categorical perception and the uncanny valley, we propose that "human-likeness" is not the only continuum that leads to negative responses when brought into relation with related categorical entities (*e.g.*, animals or robots). One example is the previously mentioned "stuffed animal," which is placed into the dimension of "human-likeness" in Mori's graph (Figure 1.3). A stuffed animal might have human-like attributes, but can still be classified as "stuffed animal". Categorical perception in scope of the uncanny valley predicts negative responses when an assignment is not clear and an entity has characteristics of two (or more) related constructs as shown for human-animal morphs, for example [82]. Difficulties in categorical perception could also explain fears and negative responses towards taxidermy animals while trying to distinguish "animal" from "stuffed animal." Thus, the uncanny valley would be then be the point of lowest familiarity due to difficulties in discriminating ambiguous lifelike entities.

Categorical perception is based on discriminating entities such as human or non-human. However, a human entity is always considered as human, even when it contains features that look not "entirely right" [145]. The perceptual mismatch hypothesis suggests that negative responses are caused by inconsistencies among different realism levels of an entity (not between different categories of entities) [192, 287]. As indicated in the related work, inconsistencies and atypical features increase the effect [187, 287]. This is supported by our second study, in which decreased realism as well as atypical features of human emotions lead to increased eeriness ratings of a virtual animal. Thus, negative responses due to high sensitivity towards deviations from typical norms and violated expectations could be caused by imperfections of humans *and* animals. This is supported by qualitative feedback in our first study using virtual animals, where missing or wrong features lead to negative responses and unpleasant associations. Our results indicate that using a uniform style (*e.g.* in R8) in a consistent level of realism leads to positive responses and is potentially responsible for the first peak in Mori's graph.

Researchers assume that the phenomenon has an evolutionary origin (*cf.* Steckenfinger and Ghazanfar [297] and MacDorman [184]). Detecting or avoiding infertile or less fit mates (*e.g.* in Neanderthals [187, 192]) cannot be explained by an uncanny valley that includes non-humanoid species. However, our results indicate that similarly to attractive and youthful characteristics of humans, aesthetic aspects of animals can potentially avoid eerie effects. This is could be shown by aesthetic ratings of stimuli R6 in Study 1, where the cat depictions with the lowest realism ratings receive significantly higher aesthetic ratings than R5. Aesthetic properties of animals allow people to draw conclusions about their well-being.

Qualitative feedback in our first study suggests that indicators of threats or infectious diseases might partially cause the uncanny effects of animals. These aspects have already been proposed as potential explanations for the uncanny valley of humans [184, 274]. Atypical fur structures, "dead eyes", or threatening poses might indicate a severe disease such as rabies, which can be transmitted to all warm-blooded species. It is conceivable that the uncanny valley is a cross-species protective mechanism for avoiding communicable diseases. This could be supported by previous research that showed reduced eye contact toward less life-like entities in both humans and monkeys [274, 297]. Such behavior generally reduces further aggression or provocative responses and helps to prevent potential threats.

We showed that theoretical frameworks such as categorical perception [32, 43, 82, 146] or perceptual mismatch [187, 192, 287] of the uncanny valley are potentially applicable and could explain uncanny effects of human and animal entities. When human-likeness is not necessarily involved as a dimension, there must be an overarching concept that also includes animals (*e.g.*, entity realism or entity category). We suggest that investigating the uncanny valley by consciously considering animals can lead to a better understanding of the phenomenon as a whole.

5.6.3 Design Implications

An uncanny valley of virtual animals would have design implications for animals in games and real-time applications. As our work shows, the realism of animal characters influences how people perceive them. We assume that this also influences how people interact with animals in games. An important issue, for example, is the question of whether an animal character should have human attributes, such as anthropomorphic characteristics or the ability to speak. This received negative ratings in our second study and is a dilemma for *e.g.* game developers who want to enable interaction with animals as with humans using speech or facial expressions. For example, players can directly talk and interact with the cat Nibbles in *Witcher III* as with a human-like character. However, our results show that the positive impression of a realistic animal is lost if it represents something other than itself. Furthermore, enabling interaction between players and virtual animals might negatively affect the acceptance of the virtual animal.

A way out of the valley would be to use highly realistic computer-graphics (RQ4). Our results indicate that native fur geometry and improved shading, correct face, body proportions, and a consonance of environment and animal can improve the visual acceptance in a way that the CG image can hardly be distinguished from a real depiction. Instead of edgy polygon models and flat textures, subdivided and tessellated 3D models using hair physics systems could improve the rendering quality of hair and fur in video games and the acceptance of human observers.

As suggested by Mori, another way to avoid potential uncanny effects is to abstract or stylize an animal character [213]. This is common practice in animated films, for example. Design implications of human- and animal-like 3D characters using stylization are well explored by previous research [133, 200, 345]. We assume that the art of stylization is well-established due to technical limitations that make it hard for designers to create artificial animals that cannot be distinguished from real animals anymore. Imperfect replicas of animals potentially cause negative feelings resulting from the uncanny valley. In our study, an abstract black and white representation caused more positive associations than current video game graphics. Furthermore, the overall impression of a virtual animal should not violate the expected nature of a real one and should not have attributes or behavior which may be considered as abnormal or threatening. In contrast, the eeriness of a virtual animal can deliberately be used to create tension or fear in horror games, for example.

In this and the previous chapters of this thesis, we used static stimuli to investigate the effects of realism of virtual characters. In the following chapter, we look into the effects of realism of animated and interactive avatars to examine their effects on the users' immersion in virtual reality.

6

Avatars in Virtual Reality

Researchers have found that the uncanny valley of the own avatar affect immersion in VR. In three studies, we show how additional factors also depend on the human realism of the avatar. We investigate how these factors are related to the uncanny valley negatively affect immersion in VR. We contribute with design implications for avatars in immersive applications and recommend to consider the user's diversity.

This chapter is based on the following publications:

V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, and N. Henze. “"These Are Not My Hands!": Effect of Gender on the Perception of Avatar Hands in Virtual Reality.” In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, CO, USA: ACM, 2017, pp. 1577–1582. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025602

V. Schwind, P. Knierim, L. Chuang, and N. Henze. “"Where's Pinky?": The Effects of a Reduced Number of Fingers in Virtual Reality.” In: *Proceedings of the 2017 CHIPLAY Conference on Computer-Human Interaction in Play*. CHI PLAY '17. Amsterdam, The Netherlands: ACM, 2017, pp. 507–515. ISBN: 978-1-4503-4898-0. DOI: 10.1145/3116595.3116596

P. Knierim, V. Schwind, A. Feit, F. Nieuwenhuizen, and N. Henze. “Physical Keyboards in Virtual Reality: Analysis of Typing Performance and Effects of Avatar Hands.” In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. Montréal, Canada: ACM, 2018. ISBN: 978-1-4503-5620-6/18/04. DOI: 10.1145/3173574.3173919

6.1 Avatars and the Uncanny Valley in Virtual Reality

Virtual reality systems transport people to other places and induce the illusion of “being there.” With the rise of HMDs, the need for understanding how and why the human brain perceives and accepts the virtual world is becoming more and more important for creating immersive VR experiences and applications. Relating to the perception of VR, authors define presence as “the sense of being in an environment” [21] or “the outcome or a direct function of immersion” [270]. VR systems and applications are usually designed to maximize the experience of presence and often make use of avatars to provide realistic and familiar interfaces between the own body and the virtual environment. First-person avatars in VR significantly increase the user’s immersion and the feeling of *presence* [149] and are particularly relevant to understand why the human brain can incorporate virtual limbs and bodies into the own body scheme.

As we have learned, the uncanny valley can cause negative associations with virtual characters and avatars. Related work shows that the uncanny valley can also occur with the own virtual body in VR [181, 298, 324]. Thus, negative associations with the own avatar cause negative associations with the overall VR experience. As the avatar is controlled with the one’s own body, we assume that a multitude of individual factors must be considered for an immersive experience of each user.

In this chapter, we explore factors related to the uncanny valley that affect the virtual experience using avatars in VR (RQ5). We look into which aspects of a virtual avatar cause negative or positive effects and interfere with the phenomenon of the uncanny valley. This knowledge will help designers and developers of immersive applications to increase the feeling of presence in VR.

6.1.1 Considering the Individual

Research in *presence* is vital, especially as games and virtual environments strive towards becoming more and more immersive. Researchers have found that experiencing presence in VR depends on a number of *individual factors* that should be considered when creating immersive applications (*cf.* Lombard and Ditton [177] and Jäger [136]). As some factors cause similar effects to the uncanny valley, it is important to differentiate between them, explicitly. For example, it has been shown that there can be gender-related differences in VR user studies: Felnhofer *et al.* [80] VR suggest that regardless of age, men generally experience higher levels of presence than women in VR. Schmidt *et al.* [266] conclude that women feel higher levels of technology-related immersion and anxiety in VR. Peck *et al.* [226] showed that the manipulation of skin tones caused changes in interpersonal attitudes and decreased participants' implicit racial bias. Not only skin color but also the virtual body size lead to biases in estimating the own weight, which was shown by Piryankova *et al.* [229]. Therefore, we assume that there are gender-related differences and individual factors in avatar perception, that must be considered while investigating the uncanny valley for avatars in VR.

6.1.2 Avatars and Body Structures

Designers of video games have unlimited freedom to vary the appearance of an avatar. Cartoonists, for example, simplify their drawings due to the thickness of black outlines. Thus, to avoid too big hands or overlapping of the black outlines, they reduce the number of fingers of their characters. This kind of stylization was adopted and preserved by many video games such as in *Earthworm Jim*, the *Rayman* series, *The Smurfs*, or *Simpsons – The Game*. Designers can also reduce the number of fingers in realistic ways. In 2009, the cover art of the game *Left4Dead 2* showed a hand with a severed little finger, ring finger, and thumb. To appease the Entertainment Software Rating Board (ESRB) game developer Valve changed the cover in a way that the index and middle fingers remained [242]. In video games, designers reduce the number of fingers in a realistic way as for the aliens in *Avatar – The Game* or Elizabeth's character in *BioShock Infinite*. Thus, the body structure of game avatars in VR does not necessarily match the structure of the user's body. However, little is known about the effects of a reduced number of fingers on the user experience and perception of presence in VR. We hypothesize that it depends on avatar realism as well as the body structure in which the users feel presence in VR.

The rubber hand illusion experiment by Botvinick and Cohen [24] demonstrated that humans can incorporate prosthetic limbs into their body representation when congruent visual and tactile feedback is provided. Further research of the rubber hand illusion (originally not situated in VR) showed how our body registers the interaction space using self-location [58], self-agency [22], and body ownership [178]. VR allows us to further explore the rubber hand illusion from a first-person view and for animated false limbs as well as full bodies [293]. The acceptance of structural changes of hands in VR was investigated by Hoyet *et al.* [125]. The authors examined the rubber hand illusion by using a six-fingered hand in VR. They found that participants experienced relatively high levels of body ownership using an additional finger when compared to using five-fingered hands. Consequently, the authors recommended investigating hands with fewer fingers. However, it has not yet been investigated how fewer limbs affect acceptance in VR (RQ6). Murray *et al.* [216] showed that VR can be used to treat the phantom pain of amputees. Not situated in VR, but also related to our work is the research by Giummarra *et al.* [97] which compared the sensations of amputees and non-amputees. Their findings indicate that both phantom pain and an illusory embodiment do not necessarily require amputation.

6.1.3 Working with Virtual Avatar Hands

While consumers currently use VR for entertainment, such as gaming, a wide range of serious applications have been proposed in the past and are currently being explored by industry and academia research. VR systems offer great potential to create pleasant working or study environments for office workers. External visual and auditory distractions can be blocked completely, which would aid productive and focused work. However, visual immersion using HMDs not only substitutes real world distractions but also makes it impossible to see the physical keyboard and mouse which are essential for high-bandwidth general purpose interaction [202].

Typewriting is one of the most used generic input methods for desktop work. To enable users to work as efficiently in a virtual environment as in a real office they require high performance input devices. Especially users who are not fluent touch typists, text input quickly becomes tiresome if they cannot see their hands or the keyboard. For efficient and immersive working there is need to provide visual cues for rendering the hand and body poses. Virtual avatars provide a familiar and natural user interface as well as visual feedback to interact with the virtual workspace. However, it is currently unknown how virtual avatar realism affects typing performance or whether it increases or decreases mental workload, especially for users with lower typing proficiency. Furthermore, it is conceivable that transparent hands can either help inexperienced typists better recognize the keys on the keyboard, or distract them while typing.

Typewriting for less proficient users is a mentally demanding task. At first glance, the arrangement of the keys is only partly understandable, so that users use the sense of sight to find the keys. Furthermore, typing with multiple fingers requires training to achieve a complicate interplay of finger placement, muscle memory, and proprioceptive abilities. This ability distinguishes many users, but is not innate. This means that the typing performance is an individual trait that everyone can learn. It also means that experienced and inexperienced typists can only be determined by a test. It is possible that this ability essentially depends on one's own appearance and can be influenced by the own virtual avatar (RQ1).

6.2 Studies Overview

In the previous section we learned that displaying the user's hand as the primary body part for interaction enables natural user interactions with the virtual world and has different effects on interaction (*cf.* Pouliquen *et al.* [234] and Azmandian *et al.* [9]). However, it is unclear whether and how individual factors and avatar realism influence the perceived sensation of presence in VR (RQ1). Mainly, it is unknown if virtual hands are perceived differently by men and women, especially when avatar hands from the opposite gender are presented.

Second, hands and fingers are an integral part of daily interaction. However, they are often removed in today's video games, for example to stylize virtual characters and to reduce their eeriness. However, whether the reduction of body parts affects the user experience in VR is currently unknown (RQ6). Investigations of the rubber hand illusion [24, 125, 293] and illusion of body ownership [7, 174, 299] are related, but different from the kind of body changes that occur when an additional limb is added. The authors of previous work highlight the importance of visual and haptic cues for registering the interaction space of the own body using additional limbs.

Third, individual differences can affect the task performance. People have different non-innate physical or proprioceptive abilities and are different in their task performance or mental workload while doing certain tasks. One challenging and important task in daily work, where different people perform differently, is typing on keyboards. Typing on keyboards is one of the most important and common input modalities in our modern life and it is likely that avatars must be considered when typing in future VR applications. However, it is currently unknown if the virtual appearance is an important visual cue for inexperienced typists. It is conceivable that experienced typists not rely on the virtual appearance and thus potentially perceive the avatar in a different way than experienced typists (RQ7).

VR systems offer great potential for entertainment, research, and workplaces. A large field of view (FoV), high visual fidelity, as well as the visual and auditory encapsulation can create truly immersive experiences with almost unlimited opportunities. Thus, the aforementioned aspects are important topics in today's interaction with avatars with the virtual world. We finally address and summarize the following research questions in three consecutive studies, situated in VR:

1. What are the effects of avatar realism and gender on perceived presence?
2. What are the effects of avatar realism and a reduced number of fingers on perceived presence?
3. What are the effects of avatar realism, transparency, and typing experience on typing performance and presence using a physical keyboard in VR?

In terms of the uncanny valley, we do not only examine how users perceive themselves when using a virtual avatar, but also other factors and how we can optimize the interaction in VR. While considering gender, we explicitly address the external appearance of our participants. This includes the individual's appearance as well as cues that distinguish between men and women. By regarding the reduction of fingers, we investigate a conscious design choice of artists or designers of immersive applications. While considering the performance in typing using avatars in VR, we examine a non-innate factor, which is not externally visible, but a proprioceptive ability that can be learned over time. By examining these aspects, we learn how these individual factors affect the VR experience and how people perceive presence. In the following, we encounter perceptual conflicts that can be resolved by careful design decisions, which will be summarized and discussed at the end of this chapter.

6.3 Study I: Realism and Gender

In our first study, we investigate the effect of different hands and gender on presence experienced by men and women in immersive virtual environments through a first-person VR experiment. Gender-related differences caused by violated expectations of how the own avatar have to look like would have consequences for designers and developers of immersive VR applications and games.

6.3.1 Study Design

The aim of our study is to investigate the effect of human and non-human virtual hands on presence perceived by men and women in VR. We used a mixed factorial design with the within-subjects variable VIRTUAL HAND and the between-subjects variable GENDER. Participants answered quantitative questionnaires in VR and provided further feedback by thinking-aloud while performing three different tasks.

6.3.2 Stimuli

We used three realistic and three artificial hand models (see Figure 6.1). *Male* and *female* hand models were taken from the Leap Motion SDK of the used sensor (Leap Motion). The *male hand* model had a haired skin texture and a muscular appearance. The *female hand* model had glossy nail textures and dainty fingers.

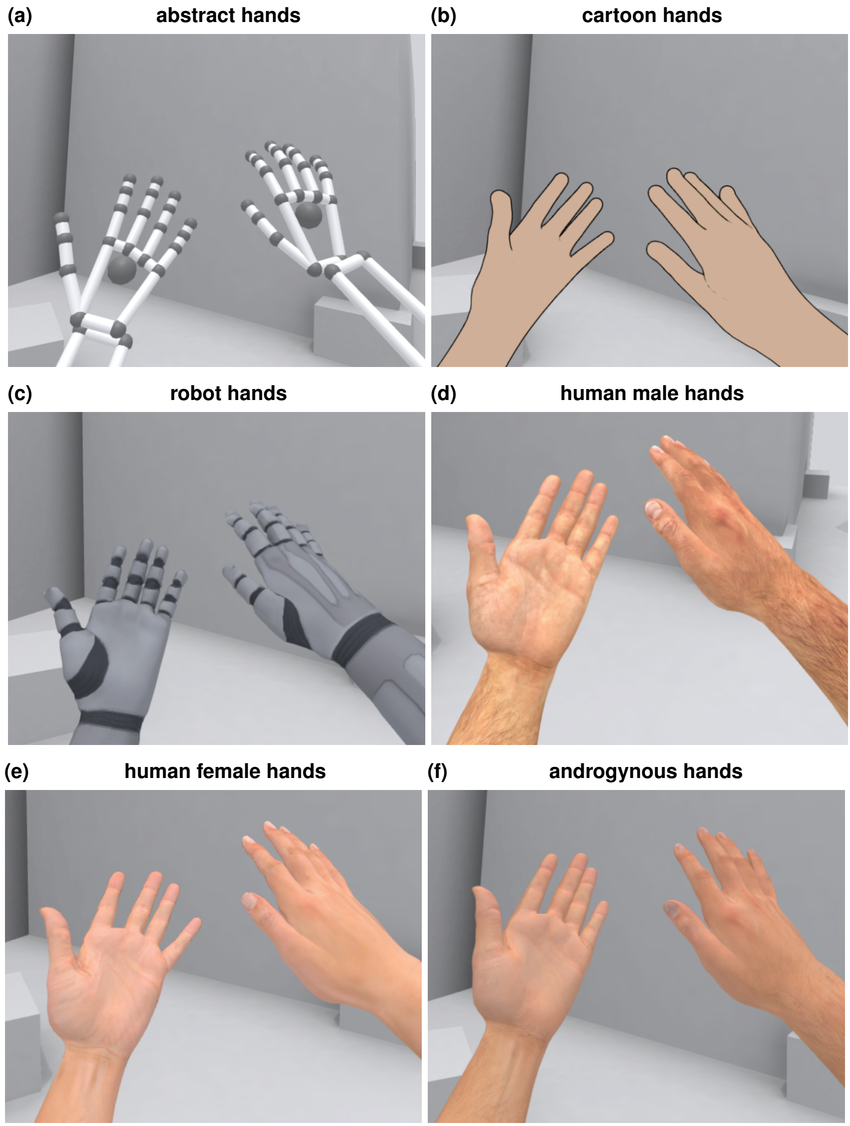

Figure 6.1: Hand stimuli in the virtual apparatus. (a) abstract, (b) cartoon, (c) robot, (d) human male, (e) human female, and (f) human androgynous hands.

An *androgynous hand* was created through an equal blending of meshes and textures from the *male* and *female hands*. We selected an *abstract*, *cartoon*, and *robot hand* as artificial hand models. The *abstract hand* model was extracted from the Leap Motion SDK and was equipped with white cylinders as bones and gray spheres as joints. For the *cartoon hand* model we smoothed the mesh topology of the *androgynous hand* and replaced the texture with a skin-colored toon shader with a black outline. The model and textures of the *robot hands* were extracted from the Genesis Bot model for the DAZ3D software application using a rigid skinning for the hand skeleton.

6.3.3 Apparatus

Our apparatus consisted of an Oculus Rift DK2 HMD, a Leap Motion sensor, and an application developed in Unity3D. Our application used the hand tracking of the Orion beta SDK provided by Leap Motion for VR support on HMDs (Figure 6.2). The Leap Motion sensor was mounted on the front of the Oculus Rift using a 3D-prined frame. Our experiment was running on a Windows PC with an Intel i7-6700, 16GB RAM, and a Nvidia GTX980. According to the refresh rate of the Oculus DK2 HMD, we set the target frame rate in Unity3D to 60 frames per second (FPS). To ensure that the FPS remains constant we designed a simple scene in Unity3D. To ensure that the tracking quality was consistent for all hands, we used throughout the same tracking system provided by Leap Motion and the same configuration of bones.

6.3.4 Tasks

We developed three different tasks which were used to ensure that the hands were present in the FoV of the participant and facilitated an immersive VR-experience: (1) in the *keyboard task* participants operated with a virtual keyboard to enter "I love VR" into a text display. (2) In the *draw task* participants painted curves and lines into the virtual space while moving their hands and performing a pinch gesture. Their task was to draw "Hello World" in 3D space. (3) In the *pyramid building task*, participants generated blocks on a virtual table by pressing a virtual button and building a small pyramid of at least 6 blocks. All scenes were blended

Figure 6.2: User in our VR apparatus: A leap motion sensor for hand tracking and an Oculus Rift DK2 HMD as virtual reality display.

using black fading. The application provides auditory feedback to confirm button presses through loudspeakers. Screenshots of the task and one panel of the questionnaire are shown in Figure 6.3.

6.3.5 Measures

Post-test questionnaires are the most frequently used measures of presence in previous work. One disadvantage of post-test questionnaires is that they rely on subjects' memories, which reflect an inconsistent and incomplete picture of the VR-experience. We therefore developed a *VR questionnaire* which appeared in front of the participant within the virtual environment. Thus, participants filled the virtual questionnaire using the virtual hands whose influence we measured (*cf.* suggestions by Frommel *et al.* [88] to avoid interruptions in immersive games). We decided to use the 32-item presence questionnaire (PQ) by Witmer & Singer [336] for the following reasons: (1) it addresses related factors as involvement, naturalness, and interface (avatar) quality, (2) the presence question-

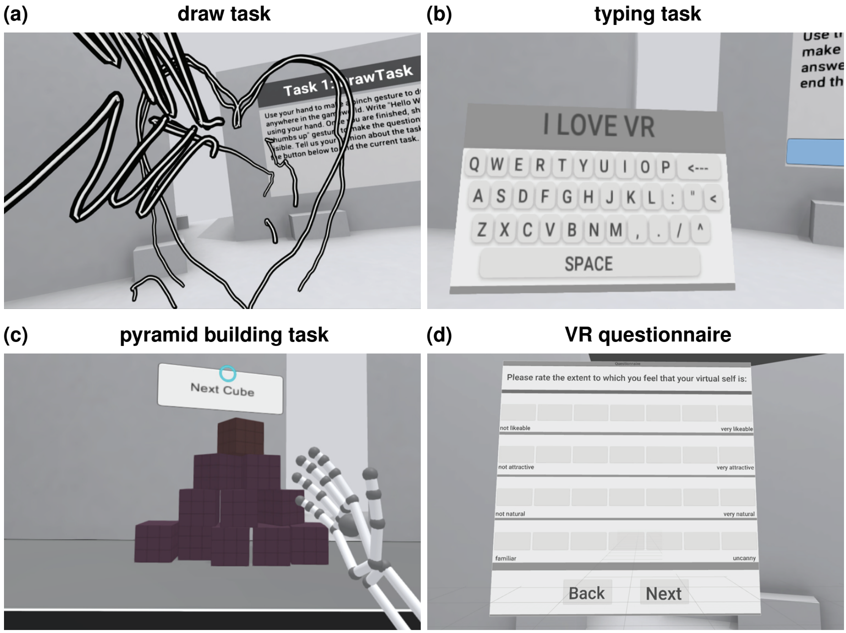

Figure 6.3: The three VR tasks and the VR-integrated questionnaire: (a) draw task, (b) pyramid building task, (c) typing task, and (d) questionnaire in VR.

naire does not include any question that cannot meaningfully be answered within the VR, and (3) the questionnaire has been used in a large number of studies (*cf.* Renaud *et al.* [243], Stanney *et al.* [296], and Youngblut and Odette [343]).

In line with our findings in Chapter 4 (p. 115) and according to our assumption that there are avatar-related factors that influence presence in VR, we asked for the perceived likeability, attractiveness, naturalness, and eeriness of the virtual hands on a 7-point Likert scale (see Section 4.2.5). We would like to note that the direct measure of perceived naturalness could be a potential confounding factor of the naturalness subscale in the presence questionnaire. Due to inconsistent or inconclusive outcomes of previous work, psychophysiological measures for presence such as heart-rate, skin conductance, or electromyography (EMG) were not used.

6.3.6 Procedure

After signing the consent form, the participant was asked to take a seat in the middle of our VR-laboratory. We presented all devices and explained the procedure of the study. The direction of the virtual space was aligned according to the speakers placed in front of the participant. After setting up the HMD, the participant was familiarized with the first virtual hand and the first task. He or she could finish the task by pressing a button or showing a thumbs-up gesture. After each task, participants had to rate how they liked the task using a 7-point Likert item on a virtual panel presented in VR. Since not all questions could be displayed at once, a virtual wall containing four questions per page was presented. The participant could navigate through the questionnaire by pressing "next" and "back" buttons. As long as participants asked for no break, they remained in VR for all hands, tasks, and questions. Using thinking-aloud, we asked them to describe their thoughts, issues, and concerns. After finishing all questions the participant repeated the procedure using the next virtual hands. After leaving the VR, we handed out a questionnaire, and asked for comments about their concerns, what they would like to improve, and which hand they finally preferred.

6.3.7 Participants

Our sample was drawn from students and employees of our university from technical as well as computer science study courses. We recruited 28 participants (14 male, 14 female) from Central Europe with light skin tones took part in our experiment. The mean age was 26.07 years ($SD = 7.99$). Students received a compensation of 10€. Seventeen participants stated that they had no previous VR experience at all; 11 had limited VR experiences.

6.3.8 Quantitative Results

On average, participants spent 58.6 minutes ($SD = 18.1$ min) in VR. Thus, the average time each participant used a virtual hand pair was 9.76 minutes. Virtual HAND and GENDER of the participants were used as factors in a multi-factorial

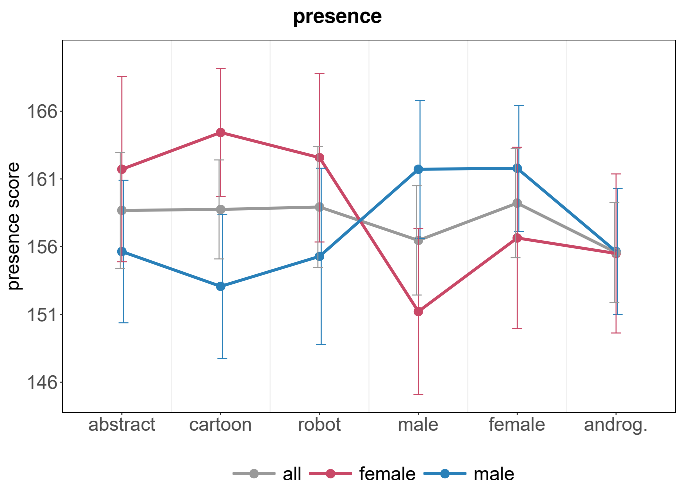

Figure 6.4: Gender related differences of perceived presence between the avatar hand models. Virtual hands are sorted by mean naturalness ratings (see Figure 6.6) of all participants. Presence scale ranges from 145 to 170. All error bars show standard error of the mean (SE).

analysis of variance of aligned rank transformed data as introduced by Wobbrock *et al.* [337]. All pairwise cross-factor comparisons are Bonferroni corrected. All means and standard deviations are shown in Table 6.1.

We found no significant effect of HANDS, $F(5,130) = .345$, $p = .884$, or GENDER, $F(1,26) = .272$, $p = .606$, on perceived presence. However, the HANDS×GENDER interaction was significant, $F(5,130) = 3.898$, $p < .001$. Pairwise cross-factor comparisons of GENDER and virtual HAND revealed significant differences between the *abstract* and *male* hand ($p = .047$), *male* and *robot* hand ($p = .034$), as well as *male* and *cartoon* hand ($p = .003$). Other pairwise comparisons of presence showed no significant differences (all $p > .05$).

HANDS had a significant effect on likeability, $F(5,130) = 5.903$, $p < .001$, but we found no significant effect for GENDER, $F(1,26) = 3.549$, $p = .071$. We found an interaction effect of HANDS×GENDER on likeability, $F(5,130) = 5.951$, $p < .001$. Pairwise cross-factor comparisons revealed significant differences

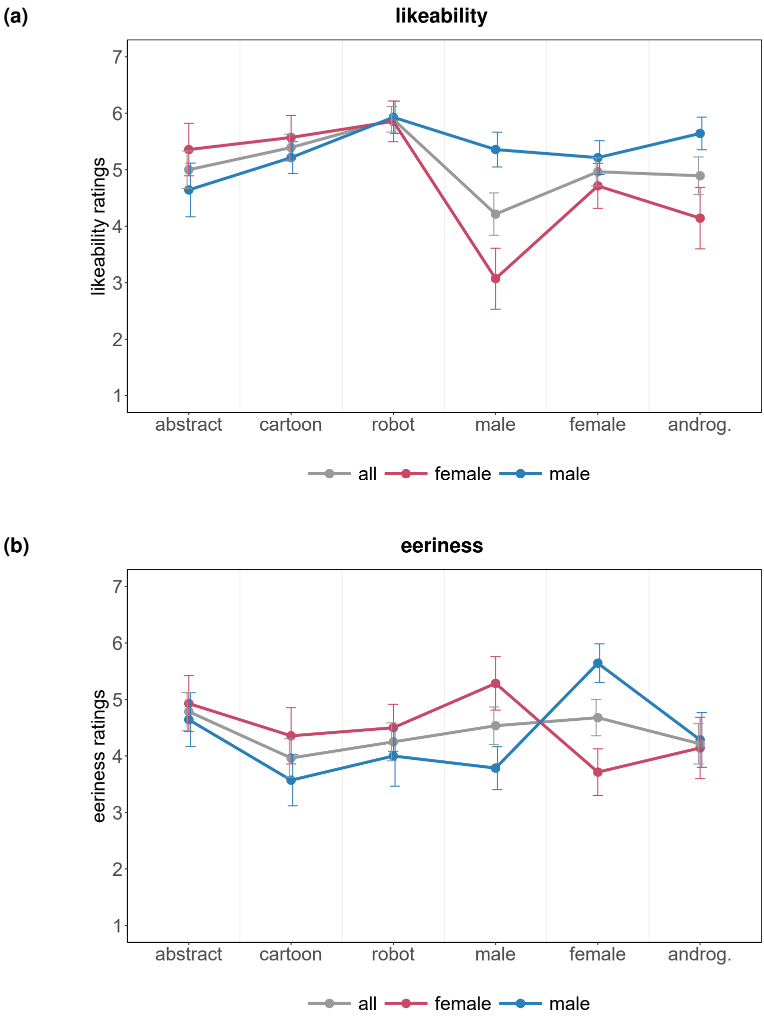

Figure 6.5: Gender related differences of likeability and eeriness between the avatar hand models. Virtual hands are sorted by mean naturalness ratings of all participants (not depicted). All error bars show standard error of the mean (SE).

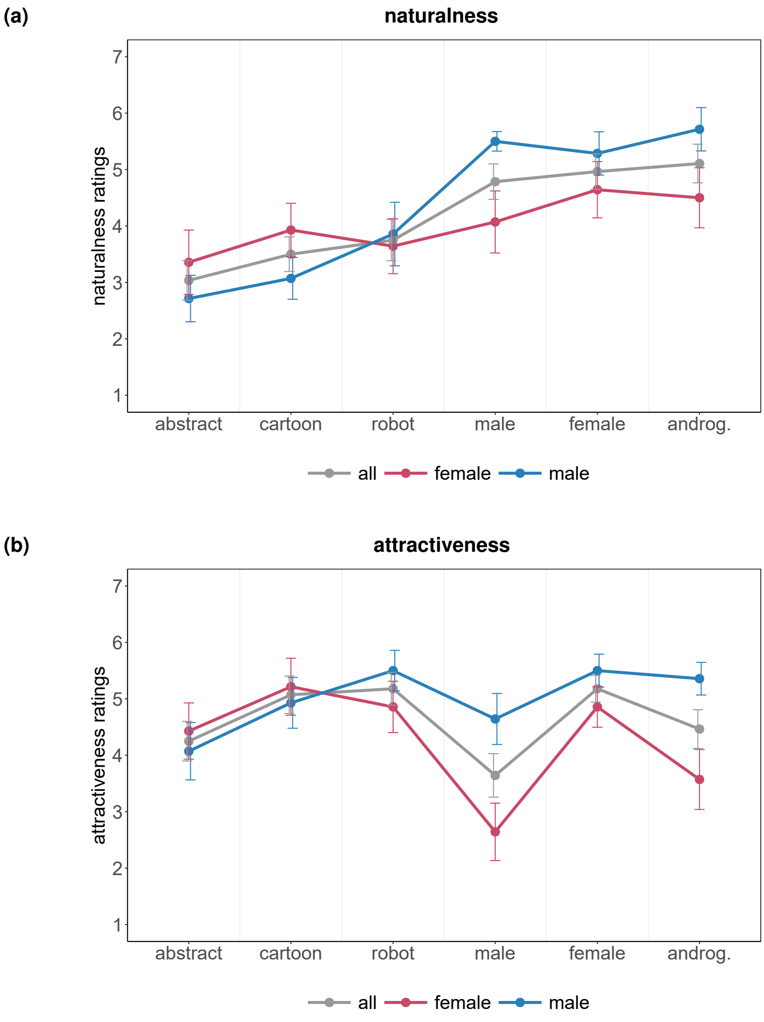

Figure 6.6: Gender related differences of naturalness and attractiveness between the avatar hand models. Virtual hands are sorted by mean naturalness ratings of all participants (not depicted). All error bars show standard error of the mean (SE).

between the following hands: *abstract* and *male* ($p < .001$), *abstract* and *androgynous* ($p = .044$), *male* and *robot* ($p = .008$), *male* and *cartoon* ($p < .001$), and *cartoon* and *androgynous hands* ($p = .043$). Other pairwise comparisons of likeability ratings were not significant (all $p \geq .178$).

We found no significant effect of HANDS, $F(5,130) = 1.181$, $p = .322$, or GENDER, $F(1,26) = .472$, $p = .498$, on ratings of eeriness. However, the HANDS $\times$ GENDER interaction was significant, $F(5,130) = 4.157$, $p < .001$, again. Pairwise cross-factor comparisons of GENDER and virtual HAND revealed significant differences between *male* and *female hands* ($p < .001$) and *female* and *cartoon hands* ($p = .011$) for the eeriness ratings. Other pairwise comparisons of eeriness ratings showed no significant differences (all $p > .05$).

To understand which avatar-related factors influence presence, we conducted a multiple linear regression using the enter method. Ratings of likeability, eeriness, naturalness, and attractiveness were used as independent variables. The regression equation was significant, $R^2 = .441$, $R^2_{Adj.} = .195$, $SE = 19.138$, $F(4,163) = 9.854$, $p < .001$, $d = .784$, for β-coefficients of naturalness ($\beta = -.159$, $p = .047$) and eerieness ($\beta = -.336$, $p < .001$). We found no significant effects on presence for likeability ($\beta = -.154$, $p = .171$) or attractiveness ($\beta = -.036$, $p = .755$). The scatterplot (not illustrated) of standardized residuals indicated that the data met the assumptions of homogeneity of variance, linearity, and homoscedasticity of the regression analysis. Assuming that the factors are independent, eeriness would explain 11% of the variance of the mean presence.

In the final questionnaire, presented after a participant had left the VR, we asked for demographics and which virtual hand they would like to use again in VR: Eight (5 male/3 female) participants would use the *robot hand* again, 7 (1 m./6 f.) the *female hand*, 5 (3 m./2 f.) the *cartoon hand*, 4 (3 m./1 f.) the *androgynous hand*, and 3 (1 m./2 f.) the *abstract hand*. One male would use the *male hand* again. Participants were also asked which hand they never would like to use again: 10 (3 m./7 f.) would never like to use the *male hand* again, 6 (all males) the *abstract hand*, 6 (3 m./3 f.) the *female hand*, 3 female the *cartoon hand*, and 3 (2 m./1 f.) the *androgynous hand*. The quantitative analysis was performed using R.

	presence				likeability			
	male		female		male		female	
HAND	M	SD	M	SD	M	SD	M	SD
Abstract	155.643	19.781	161.714	25.557	4.643	1.781	5.357	1.737
Cartoon	153.071	19.855	164.429	17.671	5.214	1.051	5.571	1.453
Robot	155.286	24.345	162.571	23.290	5.929	1.072	5.857	1.351
Male	161.714	19.052	151.214	22.871	5.357	1.151	3.071	2.018
Female	161.786	17.401	156.643	25.065	5.214	1.122	4.714	1.490
Androgyny	155.643	17.452	155.500	22.964	5.643	1.082	4.143	2.033

	eerieness				naturalness				attractivness			
	male		female		male		female		male		female	
HAND	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Abstract	4.643	1.781	4.929	.516	2.714	.083	3.357	.083	4.071	.100	4.429	.100
Cartoon	3.571	1.697	4.357	.517	3.071	.087	3.929	.087	4.929	.127	5.214	.127
Robot	4.000	2.000	4.500	.432	3.857	.080	3.643	.080	5.500	.110	4.857	.110
Male	3.786	1.424	5.286	.492	5.500	.096	4.071	.096	4.643	.061	2.643	.061
Female	5.643	1.277	3.714	.427	5.286	.111	4.643	.111	5.500	.103	4.857	.103
Androgyny	4.286	1.816	4.143	.564	5.714	.106	4.500	.106	5.357	.081	3.571	.081

Table 6.1: Means (M) and standard deviations (SD) of the virtual hand ratings and presence scores.

6.3.9 Qualitative Results

Qualitative feedback was collected through think-aloud protocols to gain deeper insights into participants' perception. Participants' comments were transcribed, annotated, and analyzed. Through open coding, we analyzed why the avatar hands affected the experience of presence. Two researchers went through all transcribed notes to check each other's coding and establish consistency in the assignment of codes to the same phenomena. A list of categories from the raw data was identified through underlining the key concepts. We found that participants' presence in VR was affected by three different deviations from their own hands:

Deviations from human appearance were mostly mentioned using the non-human hands. Participants emphasized that too abstract styles are not uncanny but not familiar enough to feel present. Participants mentioned, for example, that they feel very uncomfortable using the abstract hands: *"It's a completely unnatural arm – like a prosthesis. It distracts"* (P8, m.); and *"[...] too much abstraction feels unnatural and that abstraction is not really acceptable"* (P6, m.). One participant was confused due to the shading of the cartoon hands. *"I don't see how I should hold my hand correctly"* (P25, f). We found that female participants accept non-human hands when they consider them as *"gloves"* (P25, f.) or *"a costume"* (P20, f.).

Deviations from one's gender were perceived while using virtual human hands. Especially women reacted negatively when interacting with male hands. One female participant nearly wanted to end the VR when starting to use the male hands: *"I can't do that. This is so creepy!"* (P23, f.). In particular hair on the male arms evoke very uncomfortable feelings: *"This is so disgusting, greasy, chunky, hairy – I just want to have a shaver and wash my hands"* (P20, f.). In contrast to male participants who regarded female hands as *"very realistic"* (P13, f.) or *"unusual, but very attractive"* (P10, m.).

Deviations from the own body were noticed when participants used virtual human hands. For example, female participants drew direct comparisons with themselves when using female hands: *"Where are my freckles?"* (P19, f.), *"I hadn't French manicure! These are not my hands!"* (P22, f). *"Proportions of these hands make it clear that these are not my own."* (P8, m.). Deviations from the movement of their own hands were especially criticized by male participants. The quality of the tracking was criticized using virtual human hands, although the same tracking was used for all virtual hands: *"The tracking of this [male] hand is significantly worse. [...] The abstract hand has the best tracking"* (P2, m.).

We identified **habituation** and **excitement** (of being in VR) as additional factors with a positive effect on feeling presence while using virtual hands. Especially participants that had no VR experience were strongly involved as soon as they started a task. We observed, for example, a female participant starting with female hands and without VR experience who was immediately highly involved in the task. She naturally interacted fast and without any comments. On request, she explained that she did not even notice the hands because it *"is self-explanatory* *to interact with them in this way*." Furthermore, just being in another body was often considered positively and as exciting: *"I love it! It's fun. It's so different to be another character, another type of sense and appearance"* (P26, f.).

6.3.10 Discussion

In this work, we investigated how virtual hands influence the perception of presence in VR. Using questionnaires integrated into VR we found that presence is perceived by men and women differently. Our results show that there are significant interaction effects of gender and hands on presence. Women, in particular, feel less presence and perceive more eeriness using virtual male hands. Women feel higher levels of presence using non-human hands, in contrast to men, who feel a higher level of presence using human hands including virtual female hands. We measured lower ratings and a significant difference between ratings of female's likeability for male hands in contrast to the male's likeability for female hands. Using regression analyzes we found that the perceived eeriness and naturalness of virtual hands have a significant effect on perceived presence.

Qualitative feedback provided by think-aloud protocols reveals potential reasons for the quantitative results. We found three levels of deviations from real hands which affect the feeling of presence: Deviations from common human appearance, the own gender, or the own body. We found that deviations from the own gender were perceived negatively by female participants. Women were averse to the hair on virtual male arms and felt disgusted and were distracted. They felt comfortable when they used non-human hands and regarded them as gloves or costumes. Men highlighted the perceived realism and the hand tracking quality and felt more present with human hands.

We conclude that women have increased expectations for their representation. However, eeriness ratings show that both women and men feel discomfort when using hands of the other gender. An overall decrease of presence or likeability as predicted by the uncanny valley was not found. Lacking coherence between the perception and actual projection of the own presence as well as violations of a gender-specific appearance were potentially caused by an induced gender dysphoria – a negatively perceived mismatch of the actual (biological) sex and the perceived gender.

Due to our positive experiences with questionnaires in VR, which do not only rely on a participant's memories, we suggest to use this measuring to collect subjective data in VR for future research. The complete source code and project files with the questionnaire module for measuring the feeling of presence in VR with the PQ by Wittmer and Singer [336] is available at github.com[1].

6.4 Study II: Realism and Body Structure

Avatar realism as well as the number of fingers of the avatars' hands is an explicit design decision by artists and applications designers. In the context of games, which may reduce the number of fingers (*e.g.* for stylized hands), we decided to explore the effects on presence (*cf.* Vinayagamoorthy *et al.* [324] and Lugrin *et al.* [181]) using different avatar hands in VR.

6.4.1 Study Design

To investigate the effect of varying an avatar's number of fingers, we conducted our second study. We used a within-subject study design with the two independent variables REALISM (*abstract* and *realistic*) and FINGERS PER HAND (*two*, *three*, *four*, and *five*) resulting in eight conditions. As in our first study, data was collected quantitatively through questionnaires in VR and qualitatively through the think-aloud method and video observations. The same tasks and measures were used as in our first study (see Section 6.3).

6.4.2 Stimuli

We compared *five*-fingered hands with hands where we successively removed little finger, ring finger, and the middle finger of each hand (see Figure 6.7). We used a pinch gesture to trigger events in our apparatus and kept both thumbs as well as index fingers. We found in our first study that men and women show different levels of presence while using male and female hands (see Section 6.3). The results in the first study indicate that women dislike male hands and men perceive lower levels of presence using non-human avatar hands. Therefore, we avoided

[1] `https://github.com/valentin-schwind/selfpresence`

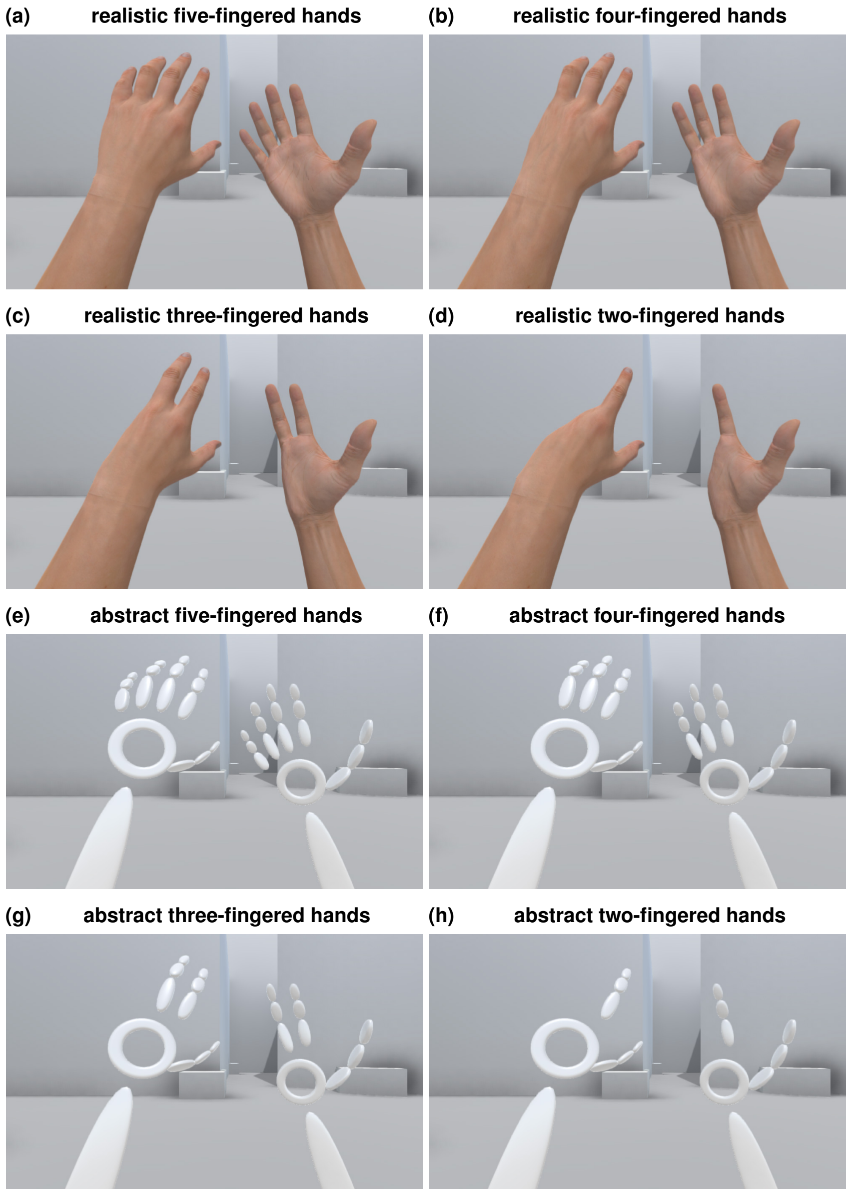

Figure 6.7: Avatar hand models with reduced fingers: Four realistic and four abstract hands with five, four, three, and two fingers.

gender swapping in VR by using the androgynous avatar model of the *realistic hand* as introduced in our first study. To remove the effect of the finger movement on the animated skin at adjacent parts of the palm, we removed the stumps of the fingers. Thus, the hands look more natural and not cut off or torn, which was achieved by smooth transitions towards the palm of the hand. A professional 3D artist removed the fingers and the influence of their virtual bones on the hand skin using Autodesk 3ds Max. The *abstract hand* models were based on the white abstract hands used by Argelaguet *et al.* [7]. They consisted of rigid oval shapes for the fingers and arms as well as a circle-shaped palm.

6.4.3 Apparatus

Our system consisted of an HTC Vive HMD and a Leap Motion sensor mounted onto the front of the display using a 3D printed frame. We used the same PC of the previous study (Windows 10, an Intel i7-6700, 16GB RAM, and a Nvidia GTX980 graphics card). Our Unity3D application used hand tracking provided by the Orion SDK of Leap Motion optimized for hand tracking on HMDs. The target frame rate of the application was 60 FPS. To ensure that the frame rate and the tracking quality were the same for all hands, we used the same tracking system provided by Leap Motion and the same configuration of bones. The surrounding scene was the same for all conditions and designed with a neutral white style and a standard sun- and skylight. Real-time global illumination, anti-aliasing, and ambient occlusion were enabled for rendering.

6.4.4 Tasks

We used the same tasks as introduced in our first study (see Section 6.3.4) and which are online available[1]. The three tasks ensured an immersive experience and that the virtual hands were in the FoV of each user. Furthermore, they enabled a versatile and immersive VR experience and could be accomplished with fewer fingers. In the *typing task*, participants operated with a virtual keyboard to enter "I love VR" into a text display. The *draw task* enabled users to perform free hand painting in the surrounding virtual space using the pinch gesture. The

[1] `https://github.com/valentin-schwind/selfpresence`

Figure 6.8: Images of the participants interacting with fewer fingers in VR situated and observed in our VR laboratory.

pyramid task was a physical simulation where participants were advised to staple together at least six blocks on a table to build a small pyramid. Black fading was used for transitions between all tasks as well as between the tasks and the final questionnaire.

6.4.5 Measures

As in our first study (see Section 6.3.5), we used questionnaires in VR to facilitate a continuously user experience. Thus, every participant filled in the questionnaire using the virtual hands whose effect we actually measured. As in our first study we decided to use the PQ by Witmer and Singer [336], which has been used in a large number of studies, includes items that address related factors such as naturalness or involvement, and all questions can be meaningfully answered in VR. As suggested by our previous studies (see Sections 4.2.5 and 6.3.5), we additionally asked for likability, human-likeness, attractiveness, and eeriness on 7-point Likert scales. Participants' feedback and their actions were recorded through think-aloud protocols and video cameras.

6.4.6 Procedure

After signing the consent form, every participant was asked to take a seat in the middle of our VR laboratory. We explained the experimental procedure and introduced the functionality of the HMD as well as the hand tracking sensor. Furthermore, all participants were instructed to "think aloud", which means to verbally articulate all their concerns and thoughts especially considering their virtual embodiments. We pointed out that participants could pause or abort the experiment at any time. The order of the eight virtual hands was given by a balanced Latin square design. The order of tasks was randomized by software. Every task could either be finished by pressing a button or showing a thumbs-up gesture. After completing all tasks a panel with the questionnaire appeared in front of the participant's view. Since all questions could not be displayed at once, panels containing four questions were presented. The participant navigated through these panels by pressing two buttons labeled "back" and "next." After leaving the VR we handed out a questionnaire where we asked for feedback, general concerns, and suggestions. Finally, we also asked them to decide both the hand they preferred and definitely not preferred.

6.4.7 Participants

We recruited 24 participants (11 males, 13 females) from our campus via mailing lists and social networks. All participants had light skin tones matching the visual appearance of the realistic virtual hand. None of the volunteers was excluded from participation in the study. The average age of the participants was 21.8 years ($SD = 6.41$). Only four mentioned having previous VR experience; 20 that they had no VR experience at all.

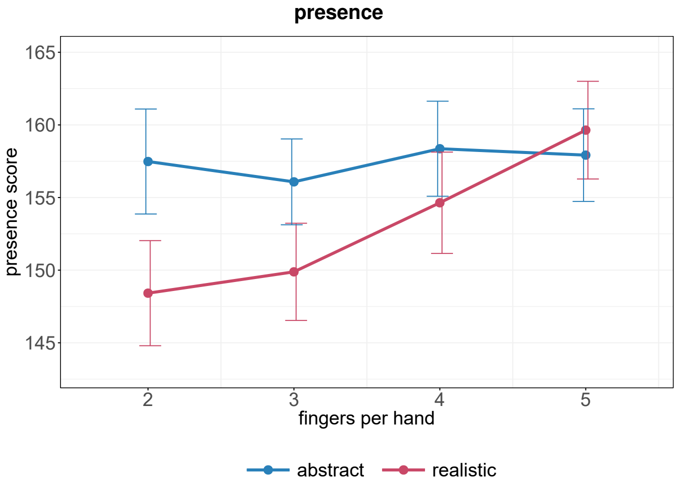

Figure 6.9: Average presence scores of abstract and realistic virtual hands with number of fingers per hand. All error bars show standard error of the mean (SE).

6.4.8 Quantitative Results

On average the study lasted for 75 minutes per participant ($SD = 8.34$). The average task completion time was 3.0 minutes ($SD = 1.8$). Two participants took a 5-minute break. We analyzed the effects of the within-subject factors REALISM and FINGERS PER HAND on our five dependent variables with analyses of variance (ANOVA) using linear mixed-effects models. All effects were taken as random at the participant level. Since we had non-parametric data, we used aligned rank transformations[1] by Wobbrock *et al.* [337]. The results of presence, likeability, eeriness, and human-likeness are depicted in Figure 6.11. All means and standard deviations are listed in Table 6.2.

Presence

A two-way ANOVA showed significant effects of REALISM, $F(1,168.00) = 13.990$, $p < .001$, and FINGERS PER HAND, $F(3,168.01) = 8.890$, $p < .001$, on perceived presence. We also found a significant interaction effect between both factors, $F(3,168.01) = 5.890$, $p < .001$. Pairwise post-hoc comparisons using Tukey's method for p-value adjustment within the levels of the main factors revealed no significant differences of the presence scores between the levels of FINGERS PER HAND using the *abstract hands* (all with $p > .05$). However, the analysis of the *realistic hands* showed significant differences between the levels of FINGERS PER HAND (all with $p < .03$), except between the *two*- and *three*-fingered ($p = .866$) as well as the *three*- and *four*-fingered hand ($p = .066$). Bonferroni-corrected pairwise cross-factor comparisons of REALISM and FINGERS PER HAND revealed significant differences between the *two* and *five*-fingered hand ($p < .001$) and the *three*- and *five*-fingered hand ($p = .021$).

Previous work found an effect of gender using male and female hands [281]. Therefore, we conducted a three-way ANOVA including the participant's GENDER as between-subject factor on presence to assess the perception of the used hands. We found no effect of GENDER and no interaction effects of GENDER on REALISM, FINGERS PER HAND, or both (all with $p > .05$). An additional analysis was conducted to determine if participants with previous experience in VR had potentially influenced the results. We found no effects of PRIOR VR EXPERIENCE as between-subject factor and no interaction effects (all with $p > .05$). The analysis of the quantitative results did not change substantially when persons with previous experience in VR were excluded from the analysis. Therefore, the data of all participants were considered in the analysis.

Likeability

For likeability, we found a significant effect of REALISM, $F(1,168.00) = 33.089$, $p < .001$, FINGERS PER HAND, $F(3,168.13) = 11.815$, $p < .001$, and an interaction effect of REALISM $\times$ FINGERS PER HAND, $F(3,168.12) = 3.603$, $p < .015$. Pairwise comparisons showed no significant differences between the *abstract*

[1] http://depts.washington.edu/madlab/proj/art/

hands (all with $p > .05$), but significant differences between the *two*- and *five*-fingered, *three*- and *five*-fingered, and *four*- and *five*-fingered *realistic hands* (all with $p < .001$). Pairwise cross-factor comparisons showed significant differences between the *two*- and *five*-fingered hand ($p = .015$).

Human-Likeness

The subjective measures of human-likeness showed significant effects of REALISM, $F(1, 168.00) = 3.956$, $p = .048$, and on FINGERS PER HAND, $F(3, 168.11) = 9.437$, $p < .001$, and an interaction effect of REALISM × FINGERS PER HAND, $F(3, 168.11) = 4.992$, $p = .002$. Pairwise comparisons showed no significant differences between the *abstract hands* (all with $p > .05$), but between all realistic hands (with $p < .001$) except for the *two*- and *three*-fingered, *two*- and *four*-fingered, and *three*- and *four*-fingered hand. Pairwise cross-factor comparisons showed significant differences between the *two*- and *five*-fingered hand ($p = .001$), the *three*- and *five*-fingered hand ($p = .05$), as well as the *four*- and *five*-fingered hand ($p = .05$).

Eeriness

For eeriness we found significant effects of REALISM, $F(1, 168.00) = 11.020$, $p < .001$, FINGERS PER HAND, $F(3, 168.14) = 17.088$, $p < .001$, and an interaction effect of REALISM × FINGER PER HAND, $F(3, 168.14) = 4.923$, $p < .001$). Pairwise comparisons showed significant differences between the *abstract hands* with *two*- and *five*-fingered hands, *three*- and *five*-fingered hands, as well as between *four*- and *five*-fingered hands (all with $p < .05$). We found significant differences between all realistic hands (with $p < .001$) except for the *two*- and *three*-fingered, *two*- and *four*-fingered, as well as the *three*- and *four*-fingered hand. Pairwise cross-factor comparisons showed significant differences between the *two*- and *five*-fingered hand ($p = .004$) and the *four*- and *five*-fingered hand ($p = .007$).

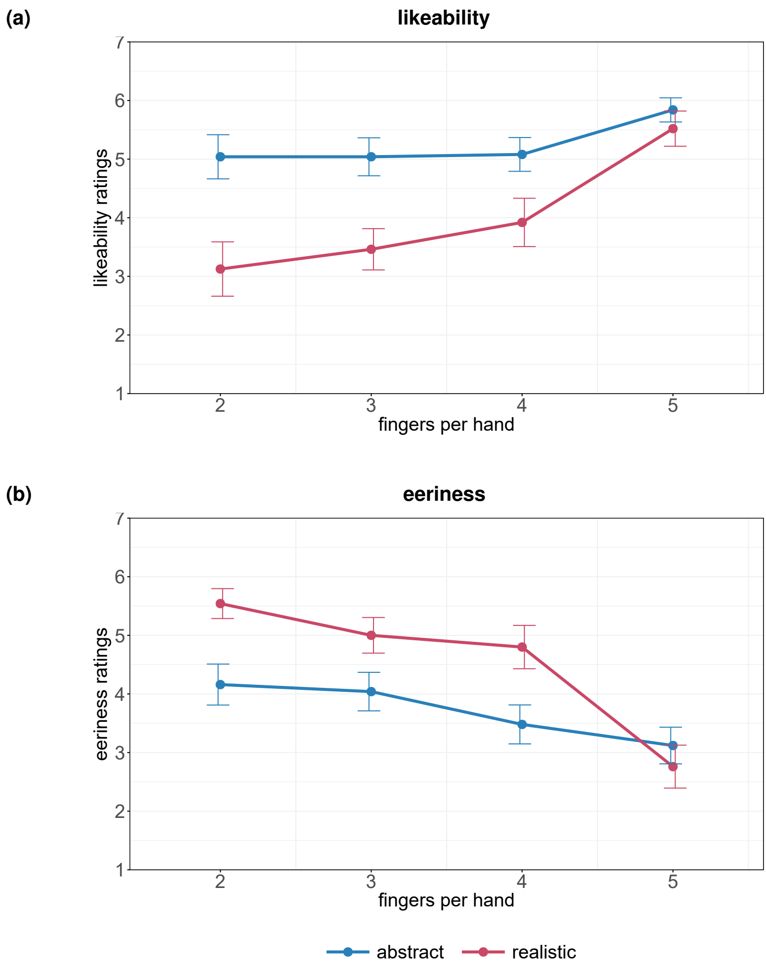

Figure 6.10: Average perceived (b) likeability and (b) eeriness ratings of abstract and realistic virtual avatar hands with number of fingers per hand. All error bars show standard error of the mean.

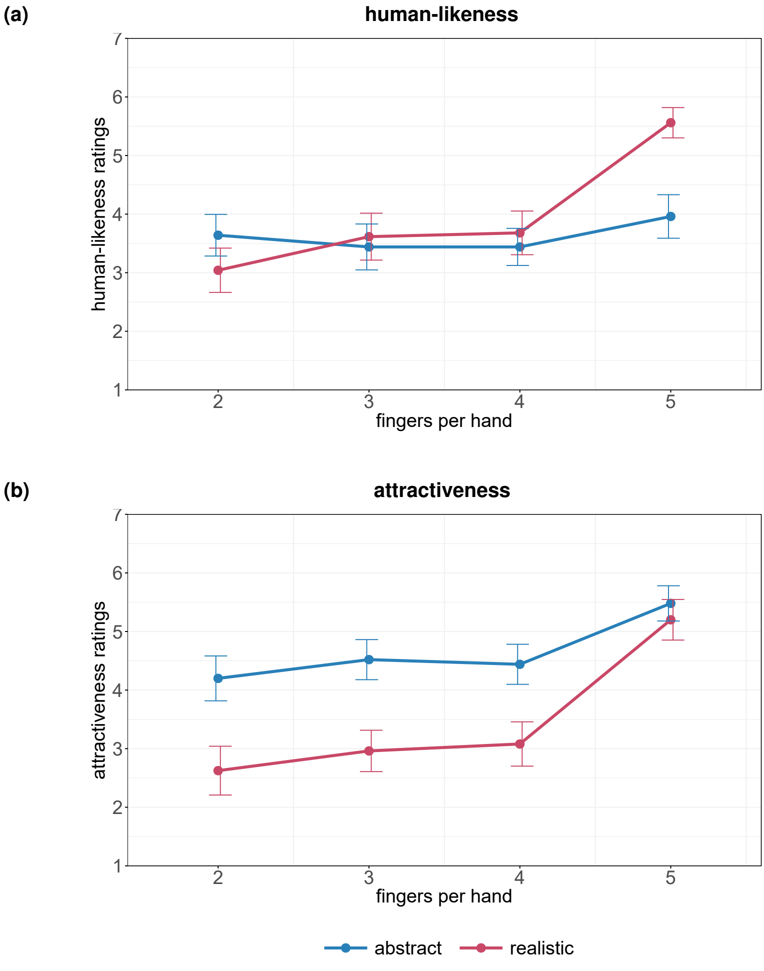

Figure 6.11: Average perceived (a) human-likeness and (b) attractiveness ratings of abstract and realistic virtual hands with number of fingers per hand. All error bars show standard error of the mean (SE).

Attractiveness

We found significant effects of REALISM, $F(1, 168.0) = 29.535$, $p < .001$, and FINGER PER HAND, $F(3, 168.1) = 19.2063$, $p < .001$, on the perceived attractiveness. There was no significant interaction effect between both factors, $F(3, 168.1) = 19.206$, $p = .09$. Pairwise comparisons showed significant differences between the *abstract hands* with *two*- and *five*-fingered hands, *three*- and *five*-fingered hands, as well as between *four*- and *five*-fingered hands (both with $p < .05$). Differences were significant between all *realistic hands* (with $p < .05$) except for the *two*- and *three*-fingered, *two*- and *four*-fingered, and *three*- and *four*-fingered hand. Due to the missing interaction effect, pairwise cross-factor comparisons showed no significant differences (all with $p > .05$).

Final Assessments

After having left the VR, a final questionnaire on a sheet of paper was handed out to the participants in which they were asked which virtual hand they most prefer and which they preferred least: Thirteen participants (52%) preferred the *abstract hand* with *five* fingers, 7 participants (28%) the *realistic hand* with *five* fingers, 2 participants (8%) the *abstract hand* with *two* fingers and 2 participants (8%) the *abstract hand* with *three* fingers. Sixteen participants (64%) would definitely not use the *realistic hand* with *two* fingers again, 3 participants (12%) the *realistic hand* with *four* fingers, 2 participants (8%) the *realistic hand* with *four* fingers, and 1 participant (4%) for each of the *realistic hand* with *five* fingers, the *abstract hand* with *three* fingers, and the *abstract hand* with *four* fingers. The outcomes are summarized in Table 6.2.

6.4.9 Qualitative Results

We collected qualitative feedback using the think-aloud method and video to gain further insights into the perceptions of our participants. Based on the records, protocols of verbal utterances and observed actions were transcribed. The transcribed protocols were annotated and scrutinized through axial coding in two iterations: In the first iteration, two researchers went through all comments to identify further individual *factors* and *effects* which had not been quantified through our ques-

RealismFingers		presence M	presence SD	likeability M	likeability SD	eeriness M	eeriness SD
Abstract	2	157.480	17.702	5.040	1.843	4.160	1.713
	3	156.080	14.475	1.843	1.587	4.040	1.612
	4	158.360	16.030	4.200	1.412	3.480	1.628
	5	157.920	15.620	1.876	1.007	3.120	1.532
Realistic	2	148.417	17.364	3.640	2.225	5.542	1.225
	3	149.885	16.714	1.741	1.758	5.000	1.518
	4	154.640	17.085	4.160	2.018	4.800	1.811
	5	159.640	16.473	1.713	1.473	2.760	1.795

RealismFingers		human-like M	human-like SD	attractiveness M	attractiveness SD	prefer N	not prefer N
Abstract	2	3.640	1.741	4.200	1.876	2	
	3	3.440	1.920	4.520	1.676	2	1
	4	3.440	1.551	4.440	1.675		1
	5	3.960	1.822	5.480	1.473	13	
Realistic	2	3.042	1.815	2.625	1.998		16
	3	3.615	2.000	2.962	1.763		2
	4	3.680	1.827	3.080	1.853		3
	5	5.560	1.267	5.200	1.697	7	1

Table 6.2: Means (M) and standard deviations (SD) of the quantitative measures in our VR study. Ratings show presence score, likeability, eeriness, human-likeness, attractiveness as well as the number of participants (N) who stated at the end of the experiment to prefer or not prefer an avatar hand.

tionnaires. One of the authors scrutinized and annotated effects; the other one factors. Both went through the results of the other and refined or complemented their results. Discrepancies between the two sets of annotations were resolved through discussion. A total of five factors and two effects were finally identified.

Additional Factors

In the qualitative analysis, we identified *association*, *habituation*, *aesthetics*, *sensitivity to display/tracking errors*, and *task performance* as non-quantified factors which influenced the individual experience of the participants in VR.

FINGERS PER HAND and REALISM were previously quantified and are not listed as individual factors in the sections below (abbreviations in the following: P# = participant ID; A/R# = abstract/realistic hand and number of fingers per hand).

Associations: We found that having fewer fingers was associated with very different prior mental concepts mainly based on the individual experience. Realistic hands with fewer fingers were associated with *"claws"* (P5, R2), hands of a *"T-Rex"* (P6, R3) *"aliens"* (P12, R3), *"mutations"* (P12, R2), or with the shape of *"pistols"* (P12, R2). The abstract hands reminded participants on characters from movies or series such as *Wall-E* (P7, A3), *I, Robot* (P7, A2), *The Simpsons* (P23, A4), or on *"crabs"* (P14, A5), *"skeletons"* (P1, A2), and *"robot hands"* (P13, A5). Associations were influenced by familiarity and prior knowledge: *"Abstract hands are much better than realistic hands because it can be that such robots have fewer fingers."* (P22, R2). Associations were also connected with the own emotional state and personal feelings: *"I feel crippled. I feel sad. When I could see my body, I would be a little crippled sad robot."* (P4, A2). Hands with fewer fingers, mainly while using abstract hands, were often considered as practical or functional tools for completing the tasks: *"These hands are more practical because I only need two fingers to complete the tasks."* (P11, A3).

Habituation: For most of the participants the study was the first VR experience. Entering VR and the impression of being in another body was at first exciting and overwhelming for them: *"Oh my god, this is so cool."* (P4, A4). The enthusiasm to be in another body outweighed a potentially strange feeling at the beginning: *"Only four fingers? Oh no! But I am so impressed. I could just look at my hands all day. That's so cool."* (P12, A4). We further observed that participants became accustomed to all virtual hands as well as to a reduced number of fingers: *"You get used to dealing with every hand very quickly."* (P22, R4).

Aesthetics: Participants were influenced by several aesthetic preferences, *e.g.* design aspects, in particular, while using abstract hands: *"Nicely designed."*

(P16, A4). We assume that design preferences are potentially connected to personal experiences and familiarity: *“To see the fingers in such a design is somehow unfamiliar. You should have put a little more effort into it.”* (P7, A3) We also found that aesthetic aspects depended on the perceived style of the virtual environment: *“Everything looks so sterile. You get used to it, however, the robot [abstract] hand fit very well into it.”* (P4, A5). Aesthetic aspects were also mentioned when using realistic hands: *“The place where the finger is missing looks disgusting.”* (P14, R4).

Sensitivity to display/tracking errors: Hands of all participants were tracked in the same way. However, some participants responded more sensitively to potential errors during hand tracking or rendering using certain hands. Then, even small errors in hand tracking were perceived as unpleasant by some participants: *“The tracking is really good unless you turn the hand around quickly.”* (P12, A4). Not only tracking errors, but the overall loss of control of one’s own body evoked negative feelings: *“I feel to have no control over my middle finger anymore. This is weird.”* (P12, R3). Some participants had problems with hand tracking that only allowed inputs when visible in the field of view of the HMD.

Task performance: We observed that participants became involved in the VR when they tried to solve a task, especially the typing task. Completing a task satisfactorily sometimes led to positive feedback related to the used hands: *“I think two fingers are even better to type or paint’.”* (P4, A2). To complete a task successfully may be influenced by the association (see beforehand) that participants consider hands with fewer fingers as useful tools: *“With those, I can type better since I’m not distracted by the other fingers.”* (P15, A2). Individual performances may be (reversely) related to other behavioral changes which are considered in the following section about individual effects.

Additional Effects

In addition to the observations described above, we identified *emotional reactions* and changes of *hand interaction* as main categories of additional or individual effects, which were not explicitly quantified through concepts in our questionnaire. We define emotional reactions as initial and prominent short-term responses. Changes of hand interaction are defined as medium-term tendencies for acting and solving problems with hands differently.

Emotional Reactions: We observed strong verbal and physical emotional reactions when participants were confronted with *realistic* virtual hands and fewer than five fingers: They felt "disgusting", "strange", "creepy", "unfamiliar", or "uncomfortable." To an extent the participants were even incensed. *"What the hell is that?"* (P6, R3). This was not the case with fewer fingers on the abstract hands: *"It doesn't disturb me that I only have three fingers because the hand is not realistic anyway"* (P3, A3). However, we recognized satisfaction of participants getting back virtual hands with five fingers after having a hand with a reduced number of fingers. *"I have my pinky again!"* (P7, A5). Some participants did not initially notice that there was a missing finger in the *four*-fingered hand condition. They were scared when they finally realized that they were having to use a *four*-fingered hand.

Hand interaction: We observed that participants changed their way of hand interaction when using a reduced number of fingers. They only used the fingers they saw. *"It is so crazy. I don't move it [the little finger] automatically"* (P4, R4). Some participants recognized without any prompting the changes of their hand interactions, which potentially led to a reverse effect on their feelings and behavior: *"It is a totally strange feeling to grab something. You don't expect to be able to hold things"* (P12, R4). They also noticed their behavioral changes after getting a *five*-fingered hand: *"Now, I move all the fingers instead of just a few, and that is more natural and immersive"* (P12, A5). Participants also tried to use haptic feedback of their real fingers to confirm that they were still there: *"Yes, I have five fingers. I see four, however, I can still feel my little finger"* (P2, R4).

6.4.10 Cognitive Mechanisms

In a second feedback analysis, two researchers repeated the analysis of the protocols provided by think-aloud. In this iteration, we used axial coding based on the identified factors and effects to understand *why* an individual factor has an effect on the user experience. In the following, we establish five *cognitive mechanisms* which potentially influence an individual's concept of having an avatar with a reduced number of fingers in VR.

Visually induced phantom pain caused by the fear of amputation and limb loss lead to strong emotional and behavioral reactions. We observed participants who painted replacements for their fingers at the stumps of their hands during the draw task. *"As if you had phantom pain. You feel it, but don't see it"* (P5, R4). We also observed that the level of associated phantom pain increased with the number of missing fingers: *"So, I can get over one finger. But not two"* (P4, A3).

The uncanny valley: Since we have only used two different styles (*abstract* and *realism*), our quantitative data does not allow us to infer the shape of a potential uncanny valley with virtual avatars and a reduced number of fingers. However, qualitative data provided while participants used *five*-fingered hands indicate that there is a potential relationship to this phenomenon: *"Looks strange. I don't know. I liked the other one [hand, A5] more because of the design. Perhaps it's because these hands look more human than the others"* (P2, R5), *"I don't know, it's confusing because the hand is too real!"* (P14, R5).

Familiarity emerges through individual prior experiences caused by associations. Associations influence personal preferences through knowledge *e.g.* about threats. Such preferences can then either be positive or negative: *"What? Please no! Reminds me somehow of claws of an animal"* (P5, R2). Familiarity influences the individual long-term habituation of using hands with fewer fingers, *"I have often seen people with a missing finger, no problem"* (P3, R4). And short-term habituation such as with one participant after getting an abstract hand with *five* fingers: *"[...] and now it's freaky to have all fingers again"* (P12, A5).

A visually induced identity dysphoria potentially causes discomfort through lacking coherence between the known appearance of one's own body and the virtually projected self. *"I feel it's supposed to be my hand. But I know it's not my hand, so I think it is creepy"* (P6, R5), The incongruence of the real and virtual body was recognized and led to a negative sensation, *"Hand is too orange"* (P22, R5), and using structural changes such as missing fingers in particular: *"What is this? I'm not a Simpson!"* (P6, R4). *"It's completely unfamiliar because you assume that you have five fingers and then you're thinking: Where's pinky?"* (P17, R4).

A mismatch of visual and haptic cues led to decreased proprioception and to a feeling of losing body control. Getting feedback from fingers that were not displayed was considered as a strange and peculiar feeling. *"It is totally creepy when I touch my fingers which are not displayed"* (P12, R2). The concept is related to the rubber hand illusion [24] when visual cues are in conflict with tactile sensations. This is also related to findings by Costantini and Haggard [54] who interpret sensory evidence about *me* is related to a prior mental representation of one's own body.

6.4.11 Discussion

In this work, we investigated how reducing the number of fingers affects the perception of virtual hands in VR. We decreased the number of fingers from the little to the middle finger and tested the hands at two different levels of realism (abstract and realistic). We collected quantitative data using questionnaires integrated in VR. Our quantitative results indicate that the number of fingers significantly affects presence and shows interaction effects with the level of realism.

The reduction of fingers did not significantly influence presence using abstract hands. However, when using realistic hands, the feeling of presence significantly decreased with the number of fingers. The diverging effect of reducing fingers for abstract and realistic hands was confirmed by significant interaction effects for all questionnaire measured except for attractiveness. Furthermore, through reducing the number of fingers, we found significant effects on likeability, eeriness, human-likeness, and attractiveness. Except for the perceived human-likeness, all measures show that the reduction of fingers lead to stronger effects while participants interacted with realistic hands. Ratings of human-likeness were constantly low for all hands except for the five-fingered human hand, which indicates that the participants had a clear concept of how human avatar hands should look like.

Through the qualitative analysis of think-aloud protocols and videos, we identified factors and effects that were not captured by the quantitative measures. We derived associations, habituation, aesthetic aspects, sensitivity to display/-tracking errors, and the individual performance (*e.g.* while completing the tasks) as additional factors that influenced the experience while using avatar hands with fewer fingers. We also found additional effects including emotional reactions and changes of hand interaction. In the second iteration of the qualitative analysis, we discussed five potential underlying cognitive mechanisms: visually induced phantom pains, familiarity based on prior experiences, the uncanny valley, a visually induced identity dysphoria, and the mismatch of visual and haptic feedback.

In the context of previous research, our work presents the first investigation of a VR experience with a reduced number of fingers. We examined the effect on presence, thus, our work contributes to a better understanding of the perception of the user's self-embodiment and avatars in VR games and applications. Our findings of the mismatch of visual and haptic cues are related to findings of investigations about the rubber hand illusion; however, adding limbs conceptually differs from removing them [24]. The illusion we created is the opposite of experiences with the rubber hand illusion and led to mainly negative feedback in VR. This is supported by findings of Hoyet *et al.* [125] who observed relatively high levels of the illusion of body ownership after adding a sixth finger. We assume that self-perception in VR using structural changes that do not match the structure of the user's body depends on whether limbs are added or removed.

Our observations indicate that the reduction of fingers induced phantom pains. The phantom pain of non-amputees could include pain due to the fear of amputation or "real" phantom pains as observed with people with missing limbs, which are also treated in VR [216]. Interestingly our participants responded emotionally, which indicated that they were highly immersed in their appearance in VR and not with the outer world anymore. Nevertheless, the feeling of presence was negatively influenced by reducing realistic fingers. Some shock moments, however, indicate that the participants still had an immersive VR experience. Source code and assets of our project are available on github[1].

6.5 Study III: Realism, Transparency, and Task Proficiency

In the previous two VR studies, we used virtual keyboards using a pinch gesture to facilitate an immersive experience. However, a virtual keyboard provides no haptic feedback and is no familiar input. The goal of the following study was to evaluate the effect of the avatar on the typing performance on physical keyboards, which were also rendered in VR.

6.5.1 Study Design

We used VIRTUAL HAND representation and task proficiency of experienced and inexperienced users (TYPING EXPERIENCE) as independent variables. To modulate the workload, we added TRANSPARENCY as independent variable to understand if additional cues are able to increase (or decrease) the typing performance for at least one of the groups. We used a mixed-factorial design with the nested within-subject variable HAND and TRANSPARENCY and the between-subject variable TYPING EXPERIENCE. For TRANSPARENCY we had three levels (*0%*, *50%*, *100%*), for the HAND factor we had five different levels (*no hand*, *finger tip*, *abstract*, *realistic*, *real*). Performance measure of the keyboard input is operationalized by words per minute (WPM). Furthermore, we investigate the overall typing experience by measuring the perceived task load and the feeling of presence in the virtual world.

6.5.2 Stimuli

Androgynous hands as used in Study I and II were used as *realistic hands*. The five-fingered non-human hand models of Study II were black shaded to increased the contrast above the keyboard and used as *abstract hands*. For the *finger tip*

[1] `https://github.com/valentin-schwind/lessfingers`

hands, we used black shaded spheres at the end of each finger. All virtual hands were rendered with 0% and 50% TRANSPARENCY. In addition, we used 100% TRANSPARENCY resulting in a *no hand* visualization and the real world scenario. An overview of all eight conditions is shown in Figure 6.12. Typing performance was measured while participants typed outside of VR on the real world apparatus, or inside of VR seeing different hands with varying transparency levels.

6.5.3 Apparatus

To investigate the different aspects of typing in a virtual environment, we used the Oculus Rift Consumer Version 1 (CV 1) HMD. We incorporated an OptiTrack W13 tracking system with eight cameras for very accurate finger and keyboard tracking. Twenty-three 4 mm retroreflective markers were affixed to anatomical landmarks of each hand to ensure precise tracking of each joint and bone of the hand. During the application startup, markers were seamlessly analyzed and automatically mapped to the virtual skeleton. In case of losing track of a marker during typing due to occlusion, our software could automatically reassigns it, when it reappeared, to untracked joints following a nearest neighbor approach. The layout of the markers is depicted in Figure 6.13.

A second generation Apple wireless keyboard was used for text input. Four retroreflective markers were attached to the top of the keyboard to enable repositioning of the keyboard during runtime to allow comfortable typing. The precise and interactive virtual replica of the keyboard was rendered according to the physical position and keypresses in the virtual environment. Our system used the OptiTrack NetNat SDK for streaming position data of bones, joints, and keyboard in real time. Our application and the virtual environment were implemented using the Unity game engine 5.4.0.

The apparatus for this study comprised two individual setups. One facilitated the real world typing task; the other allowed users to type on a physical keyboard while immersed in VR. The real world setup served as a baseline and consisted of a sixth generation 27″ Apple iMac with Intel Core i5 and a second generation Apple wireless keyboard. The computer was running a full-screen typing application showing one stimulus after another at the display. It was developed in Unity game engine 5.4.0. For the virtual reality setup, we used our developed VR system. We

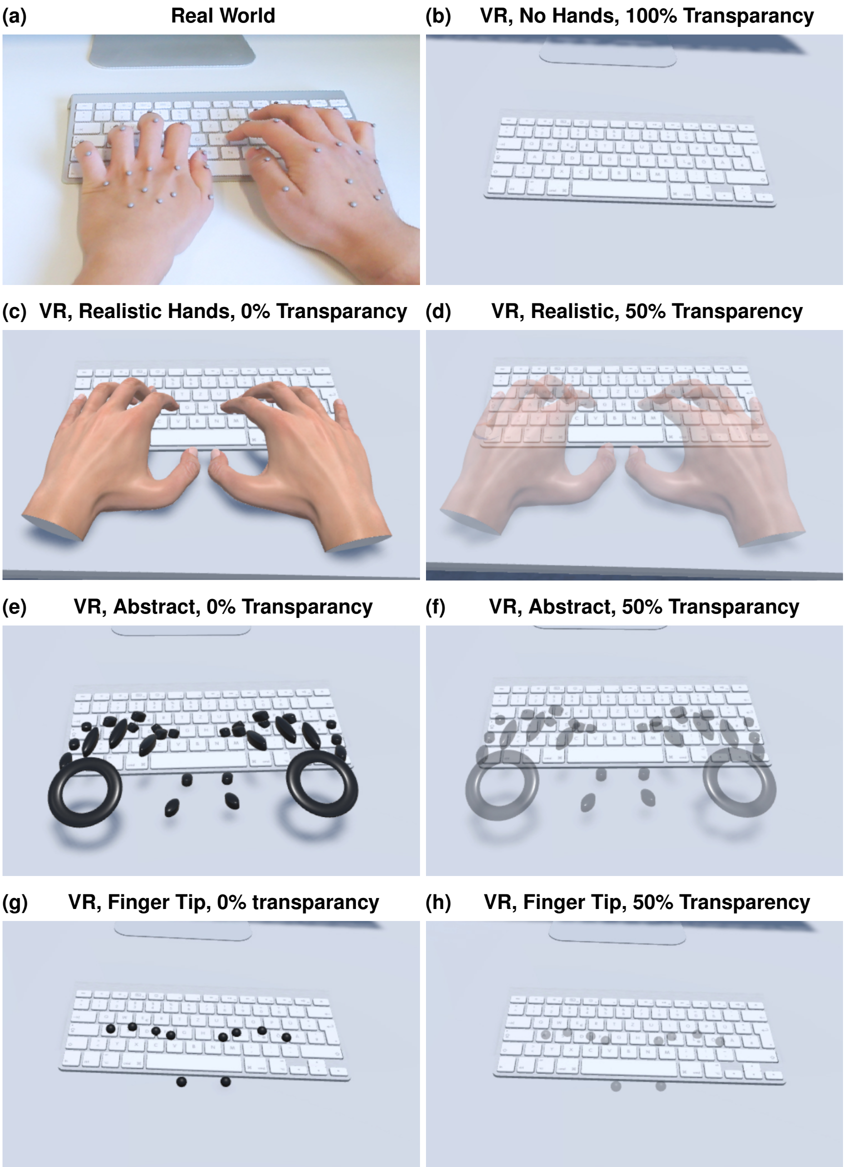

Figure 6.12: Images of the eight conditions used in the VR typing study.

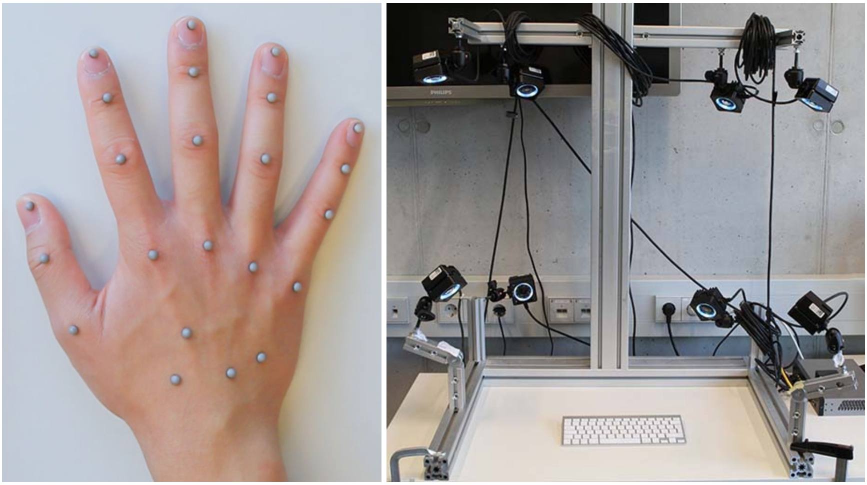

Figure 6.13: User hand and VR hardware setup. 23 passive IR reflective markers were placed on every participants' hand (left image) and the hardware setup for finger and keyboard tracking (right image).

designed an alike looking virtual environment representation of our laboratory including the real world study apparatus comprising the iMac. The real world apparatus next to the virtual replica is shown in Figure 6.14. Our experiment was running on the same PC (Intel i7-6700, 16GB RAM, and a Nvidia GTX980) as in the two previous studies. The target frame rate was set to 90 FPS to match the refresh rate of the Oculus Rift CV 1. Of course, there is a latency between a user's finger movement and photons hitting the user's retina. The calculated latency caused by motion tracking, rendering pipeline, and HMD never exceeded 30 ms during the study.

For any physical keyboard based text input users execute, they need to localize and reach out to the keyboard in a first step. Localizing could either happen visually or haptically using the surface features of the keyboard. VR HMDs prevent the user from visually localizing any physical peripherals. A system realizing effortless typing in virtual realities should support the user with an easy to understand representation of the keyboard's location in relation to their fingers. According to Feit *et al.* [77], non-touch typists' gaze switches up to 1.2 times

Figure 6.14: Side by side illustration of the real and virtual scene in our typing task. Photo of the real environment (left) and the virtual reality scene (right).

between the display and the keyboard. They spend up to 41% of their attention looking at the keyboard. Hence, an accurate representation of the keyboard and hands seems necessary particularly for this group of typists.

6.5.4 Task

In this study participants had to accomplish a simple text input task on a physical keyboard. Participants were asked to place their hands left and right next to the keyboard to mimic aperiodic typing. Being in this resting pose, a three-second countdown, displayed on the (virtual) iMac, started. After it elapsed, a random phrase from the MacKenzie and Soukoreff [193] phrase set was displayed. Participants were asked to enter the phrase as accurately and quickly as possible. Phrases were presented at the top of the (virtual) display while participants' input was shown underneath. Pressing the enter key confirmed the input and the next phrase was displayed. For each condition, participants performed three sets of ten phrases. In between each set participants had to place their hands in the resting position again and wait for the countdown to elapse. The task was the same for all conditions inside and outside of the VR.

6.5.5 Procedure & Measures

In a first step, we asked 80 participants (5 female, 75 male) to conduct a simple online typing test[1]. Based on their results ($M = 53.3$ WPM, $SD = 18.8$), we invited a random sample of 16 participants with more and 16 participants with less than 53.3 WPM to construct groups of inexperienced and experienced typists. After welcoming the participants, we asked them to sign the consent form and take a seat next to the apparatus. While attaching the 23 self-adhesive markers to each hand, we explained all devices and the course of the study to the participants. Afterward, the participant placed his hands within the tracking volume, and we defined the four markers at the dorsum of the hand as rigid bodies. In the last preparation step, we adjusted the HMD to the participant's head and calibrated it to the participant's inter pupil distance for best visual results. Then participants started with the typing task. After each task (three sets of 10 phrases), they had to fill in the unweighted version of the NASA-TLX [111] questionnaire and the PQ [336]. Subsequently, they repeated the procedure using the next hand representation. The first set of ten phrases at the start of each condition was a practice set to familiarize the participant with the different appearances. We did not include this set in our analysis. For the baseline outside of VIRTUAL REALITY, participants had to take off the HMD and move to the real setup to continue with the text input task. HANDS and VIRTUAL REALITY were presented in a counterbalanced order. Throughout the study, we logged every keystroke including the timestamp for offline analyses. After all eight iterations, we asked for comments about their experience, typing performance, and which hand representation they finally preferred. Including debriefing and detaching the self-adhesive markers, participants completed the study in 70 to 110 minutes.

[1] `https://10fastfingers.com`

6.5.6 Participants

The 32 participants (29 male, 3 female) were aged from 18 to 27 ($M = 21.9$, $SD = 2.3$). Thirteen participants had previous experience with VR. Fourteen of them were wearing corrective lenses during the study. Participants received a small gratuity and either 10 EUR or course credits as compensation for their participation.

6.5.7 Results

We conducted multiple four-way repeated measure analyses of variance (RM-ANOVA) with the within-subjects variables VIRTUAL REALITY, HAND, TRANSPARENCY, and the between-subjects variable TYPING EXPERIENCE. As previously mentioned, the within-subjects factor HAND is a nested factor of the VIRTUAL REALITY condition. TRANSPARENCY is nested into HANDS, which means that conditions of a nested factor cannot be compared with levels of factors above (*e.g.*, there is no transparency in the *real world* condition). One participant was removed from the analysis of the objective measures due to missing correct inputs (error rate: 100%) in multiple conditions. Hence, we invited one more participant from the same group of typists to compensate for the deficit. In total the participants wrote 7,680 phrases and we analyzed 5,120 phrases, since the first ten phrases of each condition were assigned for training. The results of the objective measures are shown in Figure 6.15. The mean values of all metrics are listed in Table 6.3.

Words per Minute (WPM)

The average typing performance was calculated in WPM where one word is defined to be five characters long [294]. Based on the logged keystrokes, we divided the length of the final input by the time the participant took to enter the phrase. We measured the time from the first to the confirm keypress to calculate the WPM.

We found a significant effect of VIRTUAL REALITY, $F(1,30) = 22.97$, $p < .001$, and an interaction effect of VIRTUAL REALITY $\times$ TYPING EXPERIENCE, $F(1,30) = 22.97$, $p < .001$. Furthermore, we found a significant effect of

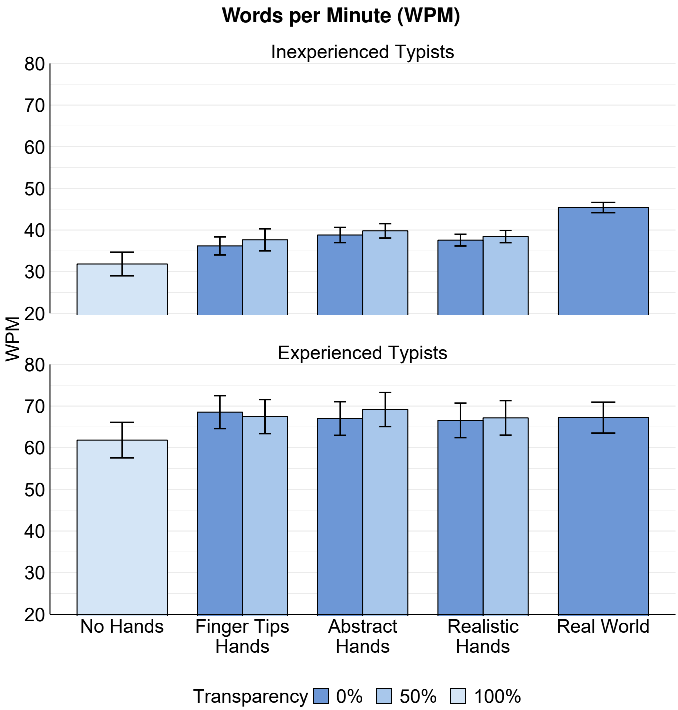

Figure 6.15: Mean values of words per minute and corrected error rate for each condition in the typing task. Error bars show standard error of the mean (SE). Exact values are also listed in Table 6.3.

HAND, $F(3,90) = 8.336$, $p < .001$, but no interaction effect of HAND × TYPING EXPERIENCE, $F(3,90) = .439$, $p < .726$. We found no significant effects of TRANSPARENCY, $F(3,90) = 1.596$, $p = .196$, and no interaction of TRANSPARENCY × TYPING EXPERIENCE, $F(3,90) = 1.022$, $p = .387$.

Post-hoc analysis was performed using Bonferroni corrected pairwise t-tests to determine statistically significant differences between the conditions. Due to the significant effects of TYPING EXPERIENCE, we compared the measures between *experienced* and *inexperienced* users separately. Due to no statistically significant effects of TRANSPARENCY, the data were aggregated across the

	VIRTUAL REALITY							
HAND	**Finger Tips**				**Abstract**			
TRANSPARENCY	**0%**		**50%**		**0%**		**50%**	
INEXPERIENCED TYPIST	M	SD	M	SD	M	SD	M	SD
WPM	36.189	8.669	37.648	1.576	38.809	7.307	39.808	6.919
Error Rate (in %)	1.244	1.528	1.304	1.127	.973	.707	1.121	.842
Corrected Error Rate (in %)	9.486	4.660	7.683	4.521	7.726	3.422	7.749	3.480
1^{st} correct Keypress (in s)	2.200	1.236	1.971	.632	1.986	.793	2.129	.842
EXPERIENCED TYPIST	M	SD	M	SD	M	SD	M	SD
WPM	68.547	15.810	37.648	1.576	67.018	16.134	69.172	16.370
Error Rate (in %)	.757	.922	.846	.974	1.003	1.135	.687	.613
Corrected Error Rate (in %)	4.354	1.858	4.766	2.297	5.118	2.307	4.889	2.909
1^{st} correct Keypress (in s)	2.108	1.332	1.831	.670	1.791	.568	1.821	.561

	VIRTUAL REALITY						REAL WORLD	
HAND	**Realistic**				**No Hand**			
TRANSPARENCY	**0%**		**100%**		**50%**		**0%**	
INEXPERIENCED TYPIST	M	SD	M	SD	M	SD	M	SD
WPM	37.581	5.611	38.430	5.839	31.848	11.338	45.398	4.909
Error Rate (in %)	1.140	1.041	1.001	1.097	.740	.538	.713	.635
Corrected Error Rate (in %)	7.681	5.031	7.712	4.313	14.015	7.549	4.904	3.388
1^{st} correct Keypress (in s)	2.456	1.290	1.864	.528	4.386	2.813	1.769	1.054
EXPERIENCED TYPIST	M	SD	M	SD	M	SD	M	SD
WPM	66.566	16.569	67.165	16.589	61.830	17.047	67.223	14.837
Error Rate (in %)	.449	.415	.362	.316	.540	.505	.597	.528
Corrected Error Rate (in %)	5.034	3.361	5.025	2.578	7.383	6.116	5.467	2.899
1^{st} correct Keypress (in s)	1.953	.754	1.980	.692	3.638	2.026	1.370	.423

Table 6.3: Means (M) and standard deviations (SD) of the typing performance indices of inexperienced and experienced typists: words per minute (WPM), error rate, corrected characters per phrase, and the time for the 1^{st} correct keypress.

transparency levels. For *inexperienced* users and the HAND factor we found significant differences between *no hands* and the *real world* condition ($p < .001$), *no hands* and *abstract hands* ($p = .024$), between *Finger Tip* and *real world* ($p = .006$), and between *realistic hands* and the *real world* condition ($p < .001$). No significant differences were found by comparing the other hand pairs (all with $p > .05$). Furthermore, we found no significant differences between the hand conditions only considering *experienced* typing users in VR (all with $p = 1$).

We summarize that the rendering of hands in VR has a significant effect on the typing performance measured using the WPM for *inexperienced* users in VR. The actual appearance of hands had no significant effect on the WPM measure of *experienced* users in typing.

Error Rate

One measure as an indicator of the users' typing performance alongside the WPM is the number of errors in the transcribed string. The *error rate* is given by the minimum string distance between the entered string and the to-be-typed phrase. It captures the minimum number of insertions, deletions, or substitutions we have to perform to change one phrase into another. We divide the distance by the length of the longer of the two strings and multiply it by 100 to obtain the error rate in percent [294].

We found a significant effect of VIRTUAL REALITY, $F(1,30) = 6.463$, $p = .016$, but no interaction effect of VIRTUAL REALITY×TYPING EXPERIENCE, $F(1,30) = 3.086$, $p = .089$ on the correction measure. There was no significant effect of HAND, $F(3,90) = 2.389$, $p < .073$ and no significant interaction of HAND×TYPING EXPERIENCE, $F(3,90) = 1.034$, $p = .381$. Both TRANSPARENCY, $F(3,90) = .158$, $p = .924$, as well as the interaction of TRANSPARENCY×TYPING EXPERIENCE, $F(3,90) = .337$, $p = .799$, were not significant. Pairwise post-hoc comparisons of the corrections showed no differences between the conditions of experienced and inexperienced users (all with $p > .05$).

Corrections

Neither WPM nor the error rate measure captures the number of corrections and edits made during text input. The *corrected error rate* [294] represents the effort put into correcting errors. We calculated the *corrected error rate* by offline analysis of the keystroke log file. Therefore, we analyzed the log file and sought characters appearing in the keystroke log file, but not in the final transcribed text.

We found a significant effect of VIRTUAL REALITY, $F(1,30) = 14.4$, $p < .001$, and an interaction effect of VIRTUAL REALITY × TYPING EXPERIENCE, $F(1,30) = 18.4$, $p < .001$ on the corrected error rate. There was a significant effect of HAND, $F(3,90) = 9.933$, $p < .001$, however, not interaction of HAND × TYPING EXPERIENCE, $F(3,90) = 2.03$, $p = .115$. Both TRANSPARENCY, $F(3,90) = 1.006$, $p = .393$, as well as the interaction of TRANSPARENCY × TYPING EXPERIENCE, $F(3,90) = 2.527$, $p = .062$, were not significant.

Pairwise post-hoc comparisons of the ratio between corrected and overall inputs considering *inexperienced* users in typing showed significant differences between all hands and the *no hands* condition (all with $p < .05$). Further pairwise comparisons considering other pairs and pairwise comparisons of experienced typists were not significant (all with $p > .05$).

Response Time to the 1st Correct Keypress

For several applications, the time to react on a specific event using keyboard input is a critical measure of typing performance. After the expiration of the countdown, we recorded the time (in *s*) a user needed for the first correct keyboard input.

VIRTUAL REALITY had a significant effect on the reaction time, $F(1,30) = 22.85$, $p < .001$, however, there was no interaction of VIRTUAL REALITY × TYPING EXPERIENCE, $F(1,30) = .19$, $p = .666$. We found a significant effect of HAND, $F(3,90) = 17.947$, $p < .001$, however, not on HAND × TYPING EXPERIENCE, $F(3,90) = .374$, $p = .772$. There were no effects of TRANSPARENCY, $F(3,87) = 1.324$, $p = .271$, or TRANSPARENCY × TYPING EXPERIENCE, $F(3,90) = .872$, $p = .459$).

Pairwise post-hoc comparisons of the average response times until the first correct keyboard input revealed significant differences between all hands and the *no hands* condition for inexperienced as well as experienced users in typing (all with $p < .001$). Other pairwise comparisons of the reaction time measure were not significant (all with $p > .05$). Thus,particularly to have *no hands* in VR affected the initial response time for the first keyboard event negatively for both *inexperienced* and *experienced* users in typing.

Further analyses were conducted to assess how the participants subjectively perceived the virtual hands. We asked for perceived work load and presence. All measures are shown in Figure 6.17.

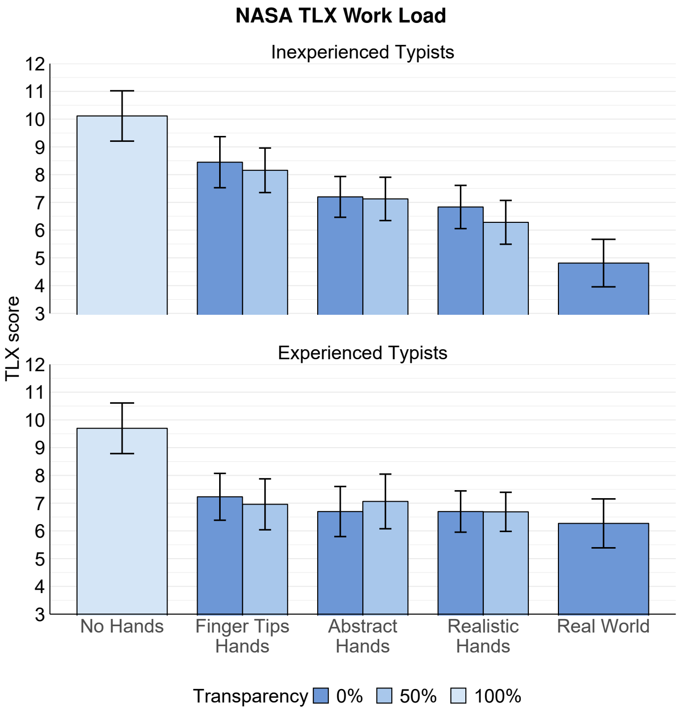

Figure 6.16: Subjective assessments of task load. Error bars show standard error of the mean.

Task Load Index (NASA-TLX)

To assess the users' perceived task load of each hand we used the score of the NASA-TLX questionnaire. We found significant main effects of VIRTUAL REALITY, $F(1,30) = 17.514$, $p < .001$, and HAND, $F(3,90) = 13.735$, $p < .001$, but no effect of TRANSPARENCY, $F(3,90) = .676$, $p = .569$. There were no interaction effects and none of the TLX measures was significantly affected by TYPING EXPERIENCE (all with $p > .05$).

Pairwise post-hoc comparisons of typing accuracy between the conditions considering the aggregated TLX measures across TRANSPARENCY and TYPING EXPERIENCE show statistically significant differences between *no hands* ($M =$

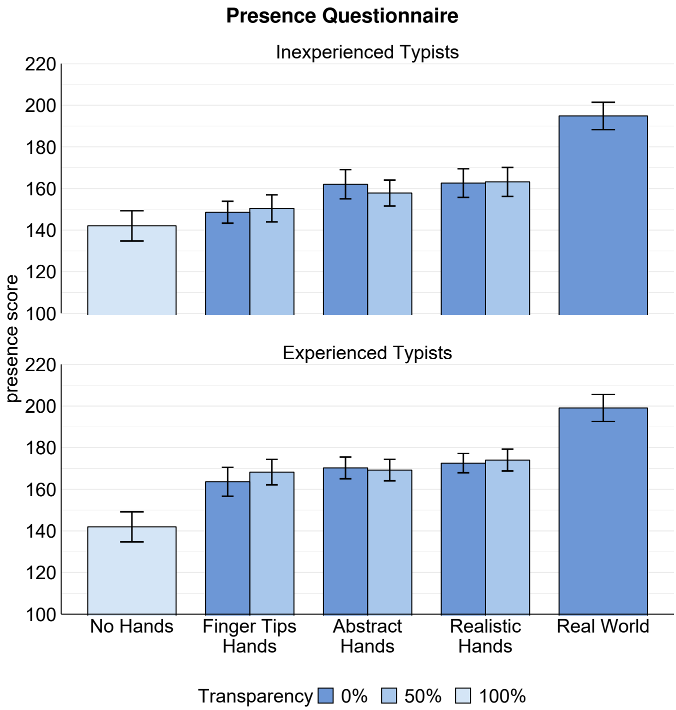

Figure 6.17: Subjective assessments of presence. Error bars show standard error of the mean.

9.906, $SD = 3.583$) and *finger tip hands* ($M = 7.698$, $SD = 3.463$, with $p = .025$), between *no hands* and *abstract hands* ($M = 7.021$, $SD = 3.348$, with $p < .001$), between *no hands* and *real hands* ($M = 5.542$, $SD = 3.500$, with $p < .001$), between *no hands* and *realistic hands* ($M = 6.625$, $SD = 2.954$, with $p < .001$), and between *finger tip hands* and *real hands* ($p = .030$).

We summarize that having *no hands* caused a significantly higher workload than the other conditions for both *experienced* as well as *inexperienced* users in typing. The lowest TLX score within the conditions of VIRTUAL REALITY was achieved by using *realistic hands*.

Presence Questionnaire

The presence questionnaire was primarily designed to compare experiences in VR [336]. For the sake of completeness and to avoid potential biases, we asked for presence in the *real world* condition as well. The overall score was averaged. Subscales are not considered in the following analysis. We found a significant effect of VIRTUAL REALITY, $F(1,30) = 99.62$, $p < .001$, and HAND, $F(3,90) = 13.269$, $p < .001$, but no effect of TRANSPARENCY, $F(3,90) = .549$, $p = .650$. There were no interaction effects and none of the PQ measures was significantly affected by TYPING EXPERIENCE (all with $p > .05$).

Pairwise post-hoc comparisons of the measures between the conditions considering the aggregated PQ scores across TRANSPARENCY and TYPING EXPERIENCE show statistically significant differences between *no hands* ($M = 142.00$, $SD = 28.49$) and *abstract hands* ($M = 164.84$, $SD = 23.80$, with $p = .002$), between *no hands* and *real hands* ($M = 196.96$, $SD = 25.78$, with $p = .001$), between *no hands* and *finger tip hands* ($M = 157.73$, $SD = 25.77$, with $p < .010$), between *no hands* and *realistic hands* ($M = 168.09$, $SD = 24.08$, with $p < .001$), between *finger tip hands* and *real hands* (with $p < .001$), between *abstract* and *real hands* (with $p < .001$), and between *real world* and *realistic hands* ($p < .001$). Other pairwise comparisons (*finger tip* and *abstract hands*, *finger tip* and *realistic hands*, *abstract* and *realistic hands*) were not significant (all with $p = 1$).

We summarize that the perceived presence was significantly affected by VIRTUAL REALITY and HANDS. The highest presence score was achieved using *realistic hands*, while *no hands* and *finger tip hands* received the lowest presence scores.

6.5.8 Discussion

Our results show that the typing performance of mainly inexperienced users using a physical keyboard in VR was significantly decreased compared to real world text input. This is confirmed by several works evaluating typing in VR [173, 202, 331]. Experienced typists' text input performances were not significantly affected by missing hands or the different hand visualizations. However, rendering virtual avatar hands significantly increases the typing performance, response time, and typing accuracy of inexperienced users. Renderings of each virtual hand pair brought their typing performance back to a level that did not significantly differ from measurements in the real world.

Our results do not confirm any effect of transparency. Related to the degree of realism or human likeness of a virtual avatar, previous work suggests an effect of the uncanny valley. As we found no effects between abstract and very realistic hands, we cannot confirm an effect of the uncanny valley on the typing performance in VR. Since the mental workload increases while typing in the virtual world, we assume that users are rather focused on the typing task than on the appearance of their hands. This finding is supported by two studies of our previous studies (see Sections 6.3 and 6.4) where participants reported that they are highly focused while performing the typing task using virtual hands and non-physical keyboards in VR. In the present study, we confirm these observations even using a physical keyboard in VR.

Our results show that the workload is statistically higher for all typists when no hands are visible. However, experienced typists' workload is not affected by typing in VR as long as hands are rendered. This leads to the assumption that hand rendering has less impact on typing performance since experienced typists do not rely as much on the visual cues. This is also supported by real world typing studies [77]. Further, realistic hands caused the lowest workload for all, while maintaining the highest presence scores for typing in VR. Abstract or absenct hands causes lower presences and a higher workload. Setting typing performance, workload, and measured presence into contrast our results suggest a correlation in particular for inexperienced typists, who seem to struggle more with abstract hand representations. We assume they need more visual guidance and abstract hands look less familiar to them. For future systems that enable typing in VR, our findings imply rendering realistic looking hands for best typing performance as well as high presence.

6.6 Discussion

6.6.1 Summary

In three studies, situated in VR, we explored the effects of avatar realism and gender (Study I, Section 6.3), fewer fingers (Study II, Section 6.4), and transparency Section 6.5 as well as typing experience (Study III).

- We found a gender related difference on the perceived presence. Women feel less presence and perceive more eeriness using virtual male hands, while men feel higher levels of presence using human realistic hands including virtual female hands. Our qualitative feedback shows that deviations of the own body negatively affect the feeling of presence.
- By reducing the number of fingers of avatars in VR, we found that changing the body structure depends on avatar realism. We discovered that additional phenomena affect the affinity as well as mechanisms of the uncanny valley in VR.
- We learned that the user's task performance influences perception in VR. Users who are less experienced in typing must depend on the representation of their hand avatar in VR. This is not the case for experienced typists, who do not rely on any representation of their VR body. Realistic hands reduce the perceived task load and increase the feeling of presence while typing. Surprisingly, rendering the avatar with 50% transparency has neither positive nor negative effects.

The main finding of our studies is that the uncanny valley is not the only phenomenon that leads to negative or eerie experiences. Beside individual preferences, we found that deviations from one's own gender, the own body, mismatched visual and haptics (losing body control), phantom pain, task and mental workload (RQ5, RQ6, RQ7) negatively affect the VR experience and the perception of the virtual body. These effects confront designers and researchers with new challenges. Users know their own appearance very well as they are familiar with their own body and we conclude that this the dimension of human-likeness in the uncanny valley in VR must be extended to the individual appearance of one's self.

6.6.2 The Uncanny Valley in Virtual Reality

Researching the uncanny valley and its effect on presence in VR is challenging, as the affinity to one's self avatar is confounded by a number of additional factors, which are neither considered in the original hypothesis of Mori [213] nor by previous work. In theory, testing the uncanny valley in virtual reality using the own avatar can be realized until the point of perfect human-likeness. However, even rendering perfect human-likeness does not consider the appearance of the own body.

In the context of the uncanny valley, gender cues and deviations as shown in our first study have significant impacts on the feeling of presence. Interestingly, perceived likeability and attractiveness showed no effect on the perceived presence. This would explain, for example, how women can perceive both high levels of presence and lower levels of realism, which is supported by Felnhofer *et al.* [79, 80]. However, Schmidt *et al.* [266] suggests that this gender related difference is caused by technology related anxieties. This is not supported by our data when we associate perceived eeriness with any occurring anxiety. Technical affinity of the surveyed participants does not support the assumption that technology related anxieties cause eerie effects using avatars.

Perceptual Mismatch of Gender-Cues

Our first study showed that additional conflicting cues (gender cues), which do not meet expectations of the own appearance, reduce affinity. This can only partly be explained by the perceptual mismatch hypothesis (see Section 2.2). In general, women have no visible hair or muscles on their arms. Thus, a perceptual mismatch due to additional cues which do not match the own appearance increases perceptual tension and causes negative affinity as predicted by Moore [210], Chattopadhyay and MacDorman [39, 187]. However, in male participants, the missing hair was not able to significantly decrease likability using female hands. Assuming both male and female VR users perceive their virtual body in a similar way, then having *lacking features* is not that worse than having *false features*. This is only partially supported by our results as lower degrees of realism do not decrease presence in females but in males.

Men show no significant changes in likability ratings. However, their perceived eeriness using female hands increases to the same degree as in women who use male hands. We conclude that the concepts of presence, likeability, and eeriness can be independent and different in VR. We assume that this is caused by further interaction effects of individual preferences and the affinity to the perceived self, which influence the overall emotional valence of the user. Higher expectations, for example, in visual realism would explain the low presence ratings and high likeability ratings of the male participants and should be considered by further research. Such aspects potentially confound or compromise perception and measurements of the uncanny valley in VR.

Body Structure and Atypical Features

In our second study, we used androgynous hands without explicit gender cues and reduced the number of fingers using abstract and realistic avatar hands. We measured the expected interaction effect between both factors and showed the relationship between the two factors. The subjective feedback provided further insights into the "somewhat obvious" result. The majority of the participants did not prefer a hand with fewer fingers (23 of 24, 95.8%). However, the majority (17 of 24, 70.8%) preferred *any* of the abstract hands more than the realistic five-fingered hand. This preference is underlined by subjective comments and in this study supports an effect of the uncanny valley in VR. However, why does the lack of fingers on realistic hands have such a strong, negative effect? We found that phantom pain, inconsistent visual-haptic integration, habituation, aesthetic preferences, negative associations, and one's own task performance influence the feeling of presence and affinity towards the own avatar. It is likely that many responses are based on a deep-rooted fear of losing one's own limbs. This fear evoke strong aversion and intense reactions to the illusion of missing limbs.

Findings in the second study can be well explained by the atypicalities in the perceptual mismatch hypothesis in uncanny valley research. Deviating features in the "human" category cause decreasing affinity in contrast to deviating features in the "non-human" category. Furthermore, in some cases we were able to induce a sensation of phantom pain. We assume that this is caused when the perceptual conflict is solved by believing that the own body is immediately being affected by limb loss. Surprisingly, the participants got quickly used to this circumstance. Either limb loss is accepted because it causes no pain or because the belief in a loss does not persist could be the subject of future research.

In Section 2.2, we discussed that research is not able to quantify the dimension of human-likeness. However, considering the body structure given by the number of fingers, we learned that this is not necessarily the case. In our study, the manipulation of human-likeness is both objectively and subjectively successful and a human characteristic, which is necessary for understanding the dimension of human-likeness in the uncanny valley framework. Missing (or additional) limbs are atypical features that are also taken into account by the perceptual mismatch hypothesis, as well as habituation and association with the own (familiar) category must cause perceptual narrowing. Therefore, we introduced missing limbs as atypical feature, which can objectively be quantified as atypical feature.

Task Performance and Transparency

In our third and final study, we learned that presence is significantly influenced by individual task proficiencies. We presented virtual hands in three different styles and congruent haptics using a physical keyboard which was rendered in VR. In the previous two studies, we observed that people were highly focused when they were typing on a virtual keyboard. By using a physical keyboard, we learned that the behavior of inexperienced typists was affected by the hand appearance. We also learned that transparency or massive changes in hand rendering did not significantly change the typing performance. Every hand, in contrast to no hands, was good enough to improve the performance of inexperienced typists. Interestingly, the relationship of hand appearance and mental workload were best in VR using the most realistic hands. Actions could therefore be carried out best using realistic hands.

There are potentially multiple answers on the question why transparency caused no effects in our typing study in VR. We assume that reduced transparency does not change the registered depth cues helping to perform the task and that 50% transparency does not provide more or less informative cues for registering one's own limb position. Based on these findings, we summarize the relation of the uncanny valley and the feeling of presence using VR avatars in a concept map which is shown in Figure 6.18. The model illustrates how the avatar in virtual reality is perceived related to the uncanny valley. Our model suggests to include the individual appearance and as one concept of realism that must be considered to investigate the uncanny valley in VR.

Previous theories of why the uncanny valley occurs include morbidity and aesthetics. Aesthetics could explain positive effects on the own appearance as they may increase the illusion of an appealing appearance and the probability to reproduce and inherit the own genes. Morbidity mechanisms prevent harmful external influences; however, the effect can also be explained by anxieties while being infected in order to prevent further spreading of a disease or to isolate oneself to avoid infections in a group.

6.6.3 Design Implications

Our findings have wide implications for designers: We recommend to avoid gender-swapping in very immersive VR applications and to provide both male and female hands if human avatars are desired. In our first study, an androgynous model without noticeable characteristics is a compromise while using human hands model. Suits, costumes, gender-neutral styles, or abstract avatar hands are rather preferred by women. The robot style seems the best trade-off between men and women for non-human avatars. This implies that customizations of the virtual self are necessary to provide optimal levels of immersion for the individual appearance and to meet the preferred preferences of each user.

Furthermore, we recommend considering the level of realism of an avatar when a reduction of fingers is desired. Using an abstract hand style in our second study, our VR users felt high levels of presence even with only two fingers left. This was not the case for realistic avatar hands. The level of presence decreased according to the number of fingers when using realistic avatar hands. Thus, when designers of games create an avatar with less fingers in an abstract way, they do not only benefit from higher acceptance in contrast to using real avatars, but also from less effort they have to spend on implementing highly realistic 3D models. Emotional reactions of our participants indicate, that a reduction of fingers of realistic hands should only be considered for shocking and horror experiences in VR. We found habituation effects which potentially indicated that users get used

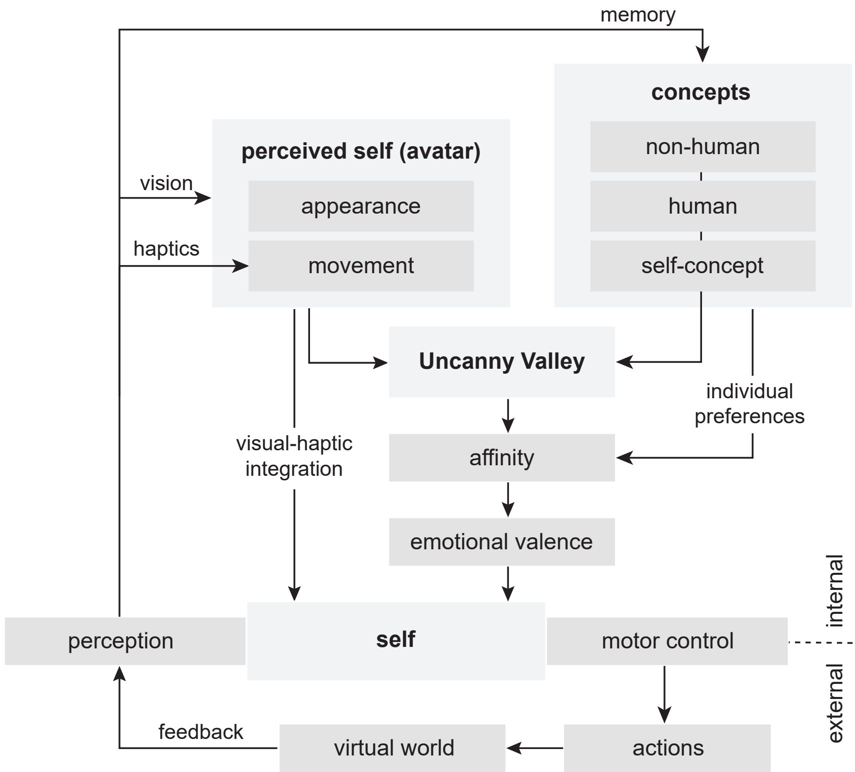

Figure 6.18: Model map of the uncanny valley occurring in VR. The model overview includes internal and external factors related to the user's perception of the own avatar. Using motor control the own body can interact with the virtual world. When the virtual world provides feedback which is perceived by the user, sensory receptors of the human body, visual and haptic cues of the virtual avatar are incorporated to the perceived self. Furthermore, sensory cues are stored in the human memory forming concepts. One concept contains the notion of *realism* including entities, humans, or the own individual. The naïve uncanny valley affects the affinity, which may be influenced by individual preferences. Affinity determines the emotional valence of the self and how we perceive *presence* in the virtual world.

to or even accept hands with a reduced number of abstract and realistic fingers. However, our participants responded sensitively to any structural changes of the avatar hands. Therefore, we recommend that the number of fingers should be kept consistent during the VR experience or game play.

In our third study, we found that virtual hands should be rendered realistically in order to target similar levels of task performance as in the real world. Experienced typists master proprioceptive processes while typing even without any hand rendering, while untrained typists rely on the visual rendering of the hand. Realistic hands reduce the perceived task load and increase the feeling of presence while typing. Thus, we recommend to design hands for physical keyboard input as realistic as possible.

6.6.4 Limitations and Future Work

In the context of our first study in VR (see Section 6.3), we suggest that deviations from the own appearance should be investigated by further research. Effects of age, ethnicity, or the own skin tone were not investigated in our study and we cannot draw conclusions about populations with other skin melanation. Findings in previous work indicate that there are ethnicity-related changes in VR [226], which could also be investigated using different avatar styles. We also want to highlight that being accustomed to the VR experience probably means that these results are only short term effects. Due to our positive experiences with questionnaires in VR, which do not only rely on a participant's memories, we suggest to use this form of measuring to collect subjective data in VR in future research.

To keep the length of our second VR study reasonable (see Section 6.4), we only used two different hand styles (realistic and abstract) . Cartoon or comic styles, which often make use of four-fingered hands for stylized characters, should be considered in further studies to examine the effect of the uncanny valley of the own avatar. We symmetrically reduced the number of fingers on both hands. Thus, we could not analyze potential effects of hand or finger asymmetries. Furthermore, we removed little fingers, ring fingers, and middle fingers. Future studies could additionally take thumb and index finger into account. Further research could also investigate the effects of other combinations of removed fingers as, for example, it is possible that the loss of the little finger is easier to handle than the loss of the index finger. With emotional reactions and effects on the hands interaction, we identified two factors in our qualitative analysis that could be quantified by future research. Furthermore, we derived five potential cognitive mechanisms that influence self-perception in VR with a reduced number of fingers.

Further investigations of typing (or working) in VR (see Section 6.5) could investigate when (or how) transparency significantly effect the input performance the user. We found no effect of 50% transparency; however, we assume that uninformative cues finally provide no additional cues for self-localization and thus, significantly decrease the typing speed of inexperienced typists as shown in the no-hand condition. Furthermore, we are interested in virtual appearance that may increase the performance. Future research should investigate whether realistic hands are the only efficient way of working in VR. It is conceivable that altered appearances potentially increase the performance of working (and collaborating) in virtual environments.

7

Conclusion

This thesis explores implications of the uncanny valley in human-computer interaction. We investigated the effect of the uncanny valley considering eyes via eye-tracking, faces using an avatar generator, and hands in virtual reality. We also investigated virtual animals to learn if the uncanny valley occurs in non-human characters. In this chapter, we present design guidelines and implications to support designers, developers, and researchers with guidelines for virtual human and animal characters and for avatar creation. We wrap up the main contributions and findings of this thesis and conclude with an overview of the this work. Furthermore, we point towards future investigations and recommendations for future virtual character designs. Finally, we close this thesis with a final remark about the uncanny valley of virtual characters and avatars.

This chapter is based on the following publication:

V. Schwind, K. Wolf, and N. Henze. "Design Guidelines for Virtual Characters to avoid the Uncanny Valley." In: *interactions* (2018). DOI: 10.1145/3236673

7.1 Summary of Contributions

In the introduction chapter of this thesis, we learned that the uncanny valley is a serious issue for designer of virtual characters. Through the increased realism in computer graphics that has come with the technological progress of computing systems the uncanny valley has become more and more important. The higher the realism of a virtual character the higher the probability that even small deviations from the human appearance can be perceived as unpleasant. This has significant impact on interaction in domains such as HCI, HRI, and CGI. The literature review in Chapter 2 provides explanations for the uncanny valley phenomenon. Previous work found that perceptual narrowing increases the sensitivity of human perception towards human faces and that multiple conflicting cues of features between non-human and human characteristics cause inconsistencies and eerie sensations. However, related work brought a number of new questions investigated in this thesis.

In our literature review in Chapter 2, we examined if there is an evolutionary origin of the uncanny valley. In Chapter 3, we investigated if altered gaze behavior as found in monkeys, also occurs in humans. We found less fixation times for characters which were perceived as eerie. We found support for the hypothesis that there is an evolutionary mechanism occurring in primates, which in turn supports the assumption that the uncanny valley is innate (RQ1). The study also supports the assumption that morbidity cues may cause eerie sensations.

For research in HCI, it is important to understand how the uncanny valley can be avoided and which tools can help to avoid it (RQ4). Previous research suggested that stylization could be a way out of the valley. In Chapter 6, we found that users of the online avatar generation *faceMaker* tend to choose physically attractive features to avoid eerie effects and that the process of avatar generation itself could be a way out of the valley. A cluster analysis of the data showed that male and female faces perceived as attractive, have similar patterns (smooth skin, natural skin color, etc.), whereas unattractive faces have multiple deviations of any kind from the human norm given by the human average. This supports the

Topic	No.	Research Question
Explanations	RQ1	**What is the explanation for the uncanny valley?** Our work as presented in Chapters 2, 3 and 5 found support for an evolutionary explanation for the uncanny valley. It is conceivable that the uncanny valley is caused by small imperfections in humans to trigger feelings of eeriness or even disgust to keep us away from potential threats or pathogen infections.
Cognition	RQ2	**Under which conditions does the uncanny valley occur?** The work presented in Chapters 4 and 5 supports the assumption that the uncanny valley occurs when multiple atypical features or inconsistencies in realism cause a mismatch of conflicting cues while perceiving entities categorically.
	RQ3	**Does the uncanny valley only apply for human-like depictions?** Our results in Chapters 4 and 5, supports the assumption that the uncanny valley exists not only for virtual human-like but also for virtual animal-like depictions. We found that paradigms of the uncanny valley that can be applied to at least one virtual animal (cat).
Applications	RQ4	**Which methods can avoid potential uncanny effects?** Perfect realism, stylization and aesthetics can help to avoid the uncanny valley for human and non-human entities (*cf.* Chapters 4 and 5). With the online avatar generator *faceMaker* we developed a tool, which can be used to learn how people avoid the uncanny valley (*cf.* Chapter 4).
	RQ5	**How does the uncanny valley of the own avatar affect immersion in VR?** Similarly to deviations from human realism, deviations from the own appearance – mainly from the own gender as shown in Study I and II in Chapter 6 – can be perceived as unpleasant and reduce affinity with the avatar at high levels of avatar realism and, thus, reduce immersion in VR.
	RQ6	**Does VR immersion with altered body structures depend on avatar realism?** The feeling of immersion with altered body structures evoke similar feelings as the uncanny valley, depends on avatar realism, and must be considered when designing avatars for virtual environments.
	RQ7	**Does VR immersion depend non-innate abilities *and* avatar realism?** We found that the feeling of presence as well as task performance in virtual reality can depend on the realism of the avatar and on the proficiency of the user to type, for example, using physical keyboards rendered in VR (*cf.* Study III in Chapter 6). This supports the assumption that non-innate factors exist that effect the affinity towards the own avatar in VR.

Table 7.1: Summary of contributions on the research questions in this thesis.

assumption that multiple conflicts must occur to cause unpleasant experiences, which support the perceptual mismatch hypothesis and further answers under which conditions the uncanny valley occurs (RQ2).

Furthermore, for research in HCI it is important to know if the uncanny valley is only related to human-like depictions (RQ3). In a quantitative and qualitative study using virtual cat renderings as presented Chapter 5, we found that two conditions in renderings of virtual animals also lead to effects as predicted by the uncanny valley for humans. We found that inconsistencies in realism and atypicalities of a virtual animal cause similar sensations as in human-like characters. This finding strengthens the perceptual mismatch hypothesis and that the process of perceptual narrowing is not only related to human-likeness, which is predicted by Moore [210]. Finally, our findings support the hypothesis that the uncanny valley is not limited to the dimension of human-likeness but to the realism of animated entities (RQ2).

In Chapter 6, we found that additional factors cause eerie effects in virtual reality using the appearance of the own avatar (RQ5). Deviations from the own gender, the own appearance, and the own body structure (RQ6) can cause very unpleasant feelings. These factors cause similarly unpleasant experiences as the uncanny valley and are related as familiar concepts may cause unpleasant and uncomfortable feelings. Considering VR the effects of the uncanny valley are intense and have significant effects on the experience as the own body is affected. Hence, we assume that the dimension of human-likeness is a special case with the similarity to one's own appearance.

It is likely that the negative experiences with one's own body in VR and with other representations are potentially triggered by the same mechanism. Uncanny effects in VR can also be explained by perceptual narrowing and a perceptual mismatch hypothesis. These effects have a significant impact on immersion and the way the user interacts with the avatar (RQ7). However, in virtual reality we must differentiate from further negative impressions. Phantom pain, mismatches between vision and haptics, the task performance and individual preferences of the user can also cause negative effects and should be considered separately when the appearance of the own body is changed.

7.2 Implications

This section gives a final overview about design guidelines for artists and developers of virtual humans, animals, and avatars in VR. The guidelines result from deductions derived from our own and existing other work, as well as from direct observations such as the *faceMaker* application.

7.2.1 Implications for Virtual Humans and Animal Characters

The uncanny valley poses serious challenges for the fields of CGI, HCI, and HRI. Improved technologies *per se* are not able to address these when artists have an unlimited number of design choices but no design guidelines to avoid the uncanny valley. On contrary, using technologies able to present highly realistic characters further aggravates the problem as even only subtle designs by artists can significantly increase the effect of the uncanny valley. In the following, we conclude our results and recapitulate how the uncanny valley can be avoided.

Avoid atypicalities at high levels of realism: One of the main findings of our work is that atypicalities (strong deviations and violations from the human norm) for high levels of realism cause negative sensations towards virtual human and animal entities. The negative effects of atypical features, such as unnaturally enlarged eyes or human emotions shown on otherwise realistically depicted animals, are larger than for characters with reduced realism. Consistently rendered realism reduces the negative effects of atypicality and increases affinity as indicated by the first peak in the uncanny valley. This means that realistic renderings and detailed textures of skin or eyes should not be combined with features that are not human-like. The negative sensations of such unnatural enlarged eyes are evident in negative reviews, for example of the character Alita (Figure 7.1) in its film trailer [175].

Avoid the "symptom of dead eyes": A virtual character's eyes are especially important when designing characters. Using eye-tracking, we found that users fixate on the eyes before they consider any other feature and assess a character as real or not real. This supports findings showing that participants rate realism and inconsistencies of humans mainly based on the realism of the eyes. This explains why skin makeup does not detract from animacy, but the eyes of a dead person do [187]. This "symptom of dead eyes" is mainly responsible for making artificial characters feel eerie and strange. Subtle features in the eyes communicate intentions, behavior, and well-being, which are all essential in the assessment of and affinity towards the depiction.

Use stylization and childish features for stylization: Research in evolutionary aesthetics states that humans respond positively to any aesthetic stimuli which are experienced to be conducive to survival and reproduction. That is why all things can be usurping and aesthetically pleasing, even if they are not realistic or humanlike (e.g. toys). Children and infants elicit protective instincts and feelings of caretaking among social communities, which are the origin of mutual understanding. Humans have these feelings even towards beings that are not of their kind, which is potentially one reason why particularly young animals, toys, and teddy bears are not in the uncanny valley. However, childish features should not be just applied to an adult as they would be atypical of this entity. Features of entities that are known from childhood and positively associated can also avoid this effect and can deliberately be used to avoid the uncanny valley.

Use aesthetics and appealing features: Designers can increase characters' physical attractiveness to avoid uncanny effects. While creating faces using our faceMaker avatar generator, participants used appealing features when designing human faces: Slight asymmetry and smooth skin caused by reduced surface details are the most important factors in physical attractiveness. Healthy and natural skin tones, natural and realistic proportions, clear gender cues, and volumetric and healthy colored lips are further important features of attractive virtual faces. Interestingly, perfect faces are not desired either. Most participants reduced details on the skin, but not completely. This supports the assumption of previous work saying that slight imperfections are still more appealing than perfectly smooth skin or facial symmetries. Regarding skin realism in face design, there is a narrow burr in modulating the desired affinity. On the one hand, perfection is not as appealing as slight imperfection; on the other, too many imperfections cause uncanny effects. This is a paradox and a hard challenge for designers aiming to increase the likeability of their characters. However, as foundation for creating faces, we can highly recommend the average face. It is the ideal base mesh to make faces appealing and has the least distance to all human faces, which significantly simplifies customization and individualization for the user or artist.

Use the creativity of the users: For our research the avatar generator faceMaker helps us to determine the preferred characteristics of virtual faces and to avoid unpleasant effects of the uncanny valley. For interactive applications in HCI, such as games, an avatar creator allows individual customizations for every user. Our system offers 37 parameters including the most important features such as gender, skin details, hair color, lips volume or skin color. Interestingly, if having a choice, users are more willing to create human-like instead of cartoon-like faces and bridge the valley using very attractive and individual features. Furthermore, users put a lot of effort into expressin<g themselves by individualization of an avatar. Thus, we learned that avatar customization itself can help users create the avatar they like and to identify themselves with it, which decreases the uncanny valley. Considering games, customizing the own appearance could not only enhance the social presence [65] but also enhance the replayability in games using a new character, for example.

Consider the effect in animals: The uncanny valley predicts that further reducing the degree of human-likeness avoids the uncanny valley. Our findings clearly show that this is not necessarily correct. Animals are by nature less human-like, but can also fall into the uncanny valley. The designer must instead be aware of the category of the character and either avoid resembling a realistic, existing category – or make it indistinguishable from real humans. Our results indicate that stylization or abstraction can mostly be done by significantly less realism of the skin, fur, or eyes.

Use the uncanny valley effect: Designers of virtual characters can use the uncanny valley for villains and ambiguous characters to deliberately elicit eerie effects. One example of such a villain is Gollum from *Lord of the*

Figure 7.1: Image from the "Alita: Battle Angel" Trailer. ©2017 by Twentieth Century Fox, Quelle: Twentieth Century Fox/medianetworx press service, released for print and online.

Rings, where the transformation of his former human appearance was even a subject of the movie. His atypical appearance is accepted as it underlines the nature of the character. Villains such as Peter Cushing's CGI double of Tarkin in *Rogue One* [163] or Jeff Bridges' young double in *Tron 2* [62] are more controversially discussed as their virtual clones try to resemble real actors. However, as introduced at the beginning, Rogue One was not mainly criticized for showing the double of a villain but for the short appearance of Carrie Fisher's young double of Princess Leia at the very end of the film.

7.2.2 Implications for Virtual Reality

In VR, the effects of the uncanny valley are much larger than in other modalities. The fusion of vision and tactile sensations dramatically increases the experience of being someone or somewhere else. If the user is supposed to feel immersed with the own avatar in VR, multiple aspects should be considered to avoid negative effects of the uncanny valley:

Consider the user's diversity: Users are very familiar with the characteristics of their own body. If the user's virtual body is rendered with a high level of realism, even small deviations can cause strong discomfort. This is especially apparent when women use a male body and see hair on their arms. Not only the gender but also skin color and special features should be considered to enable immersion. Alternatively, less realistic, androgynous bodies without specific gender cues or virtual clothes can help to avoid unpleasant effects in VR. A camera mounted on a VR headset, for example, could help to render the user's real body in VR. However, this can reduce immersion due to inconsistent styles and lighting between the real and the virtual world. From our experience with avatar generators, customizations of the virtual self also help to avoid unpleasant effects in VR. The ability for users to customize their appearance according to their preferences could help them to feel more immersed.

Avoid altered body structures at high levels of human realism: Designers of stylized characters often change the physical structure of their characters. Cartoonists, for example, simplify their drawings and reduce the number of fingers. This kind of stylization is adopted and preserved in many video games. We found that a reduced number of fingers does not affect the feeling of presence with an abstract avatar [5], as users consider their avatar as a useful tool or body extension and accept altered body structures using abstract styles. However, designers should not reduce the number of fingers at high levels of realism. Reducing the number of fingers of realistic human hands in VR causes very unpleasant feelings and has a negative effect on immersion. Some of our participants felt phantom pain, a strong disconnection to their own body, and confusion from the lacking congruency of vision and haptics. Thus, high levels of human-likeness imply a matching body structure for immersive applications using virtual avatars.

Users may get used to their avatar: Users can quickly get used to their avatar and the VR experience. In our studies, we observed that participants stopped complaining about their appearance after only a few minutes. Some mentioned that after they got used to a virtual appearance, they felt disturbed when their body changed again. Thus, we assume that habituation effects may occur and users will learn to interact with their avatar even when deviations and inconsistencies with the real appearance exist.

Consider depth cues while stylizing: The task performance of users can be affected by the avatar when no depth cues are provided. Investigating typing in VR, we found that users with different typing proficiencies profit differently from having more realistic avatars; inexperienced typists look more at their avatar hands to coordinate their movements above the keyboard. We assume that inexperienced typists need more depth cues in contrast to experienced typists, who do not need to refer to depth cues to localize their hands. Additional surface structures, shadings, shadows, and lights support of the binocular disparity and provide more control for localizing the own body movement. Thus, designers and developers should consider if the users can smoothly perform such tasks using the virtual hand within their application. This should be considered when, for example, hands in a 2D cartoon style are desired.

Avoid visual-haptic mismatches: Mismatches between the virtual and the own appearance, such as removed fingers, were confusing, and perceived very negatively. Visually missing limbs, which still perceive physical haptic feedback cause a strange feeling and unpleasant feedback. If fewer fingers are provided for avatars in VR, no tactile or force feedback should be presented, to avoid mismatches of vision and haptics.

Extending the uncanny valley: Multiple studies show that the uncanny valley is not limited to the appearance of humans or the dimension of human-likeness at all [39, 187, 279, 281, 285]. Both animals and the own appearance in VR are beyond the scope of the dimension of human-likeness. Thus, the problems of operationalizability as discussed in Chapters 1 and 2 continue to exist. So, to which continuum does the uncanny valley actually refer? From the point of view of computer graphics, the answer is simple and obvious: virtual realism. Nevertheless, robots and dolls in the real world are affected by the uncanny valley, too – and they are “real.”

Related work indicates that perceptual narrowing is a process that happens in animals and humans, but obviously not in all objects [39, 187]. The example of leaf spots [192] potentially evoking less disgust in us than expired food or moldy rooms is mentioned by Moore [210]. If these examples evoke negative feelings due to the same evolutionary mechanism, the uncanny valley would be an important tool, evoking repulsive feelings in us to point out abnormalities in any things or sentient beings that could threaten our health and existence. In that case, the uncanny valley would have been coincidentally referenced to humans – but is in reality a natural instinct of keeping distance from things that seem familiar but can pose a threat to health.

However, why does lacking realism in CG renderings trigger this feeling? Artists, who dare to depict virtual humans despite the effort of making them, take synthetic materials (graphics shaders; lights) and assets from the real world (photo textures; model scans) and combine them in character creation. The results are inconsistencies that occur in various areas (skin, mouth, eyes) and lead demonstrably to the uncanny valley effect [39, 187]. So the effect occurs when some regions do not look like others – similar to the appearance of infections of a body or object with bacteria, molds, or viruses. We strongly recommend that future work investigates these aspects and that artists render virtual characters consistently at any point.

7.3 Future Work

This thesis is also aimed to provide common ground for future work in the fields of uncanny valley research in HCI, CGI, games, and VR. In the course of this work we discovered new challenges beyond the scope of this thesis. In the following, we point to directions for future research.

Eye-tracking and the uncanny valley: We suggest that future research should investigate which kind of fear human-like entities provoke after first eye contact and how short and long-term interaction with robots or avatars changes. A further question is whether abnormality, as well as ambiguity, arouse the same eerie feeling after eye contact. In particular the use of animated and interacting artificial characters should be the topic of further research using eye-tracking methods.

Further investigations using *faceMaker*: In Chapter 4, we used likeability and physical attractiveness as indicators to learn more about features that make a virtual face more likeable or appealing. This could be extended using further indicators of interest. Shares of more realistic attributes (e.g. freckles, moles, acne, wrinkles, beards, hair, more cartoon styles) that affect the impression of skin also should be further investigated. With faceMaker researchers can collect more data within a longer period of time to investigate gender related differences, cultural preferences, and possible differences between groups of ages. Technical limitations include the lack of volumetric hair and fine geometrical displacements instead of bump maps. Other questions which have arisen are whether people perceive a face differently from the user who created it, and how they would rate faces they already know or that look like themselves.

Investigations of virtual animals: The studies in Chapter 5 only consider virtual renderings of cats. It is possible that the uncanny valley affects other virtual animals and artificial animals in the real world (e.g. pet robots, stuffed animals) in a similar way. We only assume that our findings are also valid for other animal species and strongly suggest future research to explore further animal species. Furthermore, there are complex interactions between humans and domesticated animals, which are not currently not regarded. Breeding, for example, has changed the appearance of animals to make those breeds more appealing to humans. This could be considered in further studies comparing different kinds of animals and breeds. Here, a direct comparison between familiar (e.g. pets) and less familiar animal species could be investigated as well. Future work should also pay attention to familiarity and positive associations or individual prejudices against certain animals. Studies on discriminating humans and animals indicate that there are different thresholds for categorization human from non-human and animal from non-animal entities [32, 38]. This should be considered in developing questionnaires aiming to quantify human- and animal-likeness with respect to the uncanny valley. Future work could also further differentiate the feeling of uncanniness from other feelings such as fear, disgust, or horror.

Methods for quantifying interaction with virtual avatars: Effects in human-avatar interaction with animated characters are difficult to quantify in experimental research. We suggest to develop novel methodological approaches to measure impacts of interaction with virtual characters. The social quiz game "Who Am I?", for example, could help to simulate natural human interaction with avatars remotely using a controlled Wizard-of-Oz setup. With low-cost hardware such as a Microsoft Kinect v2 and an eye-tracker, it could be possible to capture emotional expressions, speech, and mimicked actions of a real person. Animations can be transferred to the responses remotely. Thus, animated interaction could quantify the experience using reliable measures (the responses) in a true human-avatar interaction scenario.

Effect of the uncanny valley on visual-haptic integration: In the third study in Chapter 6, we learned that visual and haptic cues of the avatar are incorporated into the perceived self in VR. Hand tracking and haptics are gaining more importance as key technologies of VR systems. For designing such systems, it is fundamental to understand how the appearance of the virtual hands influences user experience and how the human brain integrates vision and haptics. However, it is currently unknown whether multi-sensory integration of visual and haptic feedback can be influenced by the appearance of virtual hands in VR and if the uncanny valley modulates this influence. We recommend to conduct a user study to gain insight into the effect of hand appearance on how the brain combines visual and haptic signals using a cue-conflict paradigm. The question is, if sensitivity changes in sensations of tactile or force feedback correlate with the degree of how the brain incorporate the virtual avatar into the own body scheme. These findings would help to understand if humans incorporate the appearance of their own body in haptic sensation and if VR technology is able to manipulate the visuo-haptic perception.

Avatar realism and gestures in VR: People can use VR to collaborate together, even when they are actually in different locations. To collaborate, they can use avatars and gestures, and the most important human expression in non-verbal communication which is pointing. This gesture is not only used to visualize directions but also to interact with computing systems especially in VR. Previous work has shown that determining the precise pointing direction is affected by inaccuracies of human proprioception and multiple methods have been proposed to determine the correct pointing direction [197]. Furthermore, it is known that there is a difference between gestures in the real world and in VR [198]. However, it is currently unknown how people perform pointing gestures in VR when the own body is either invisible or has a completely different appearance than the own one. The question arises if the uncanny valley or avatar realism has an effect on the accuracy of pointing gestures in VR. Which cues cause potentially irritating or distracting visual cues? This aspect could be important to understand how the own avatar changes the expression of gestures and interactions.

7.4 Final Remarks

While I worked on this thesis, three questions were generally asked by other researchers, colleagues, friends, and family. The following two questions were the most frequently asked ones:

"What is the uncanny valley?" and *"What causes the uncanny valley?"*

My default answer is that the uncanny valley is a phenomenon in Psychology that hinders virtual characters or robots from being accepted when they are almost – but only almost – look like real humans. Why the phenomenon exists is unclear, however, some researchers assume that the uncanny valley prevents us from threats such as infections or strange behavior. My research is dedicated to the impact of the uncanny valley in human-computer interaction as this discipline is mainly affected by the uncanny valley where developers, artists, and researchers use virtual characters, avatars, or robots.

After I described what the uncanny valley is, many people remembered their last eerie encounter with artificial figures. For example, some describe their strange experience in *Rogue One: A Star Wars Story* (2016) [70] when General Tarkin or Princess Leia appeared on the screen. Although Leia was only visible for some seconds, people felt that it would ruin at the very last moment at the end of the film. When Tarkin appeared, the spine-tingling feeling when observing the villain was somehow okay, because "actually he should be scary." In this moment, people realize that the uncanny valley exists somehow and that it is affecting their life and opinion in some way. In research of this nature, the "researcher's dilemma" comes into play. When people become aware of the uncanny valley and its close-to-subliminal but very real effects they will look more closely and pay more attention to their feelings the next time they see a CG character, compromising a more natural response. Some people may then even suspect an uncanny valley in all sorts of things such as in physical simulations [35] or after they date people after a long relationship[1]. In other words, the awareness of the uncanny valley affects peoples' responses and biases.

However, this could be a good thing. The uncanny valley is now becoming part of human culture and Mori gave people a name to express an intuition or suspicion which nobody before could really express. Yes, today we may see the uncanny valley in all kinds of things eluding a general prediction about their "okayness." However, now I know that it is more important (in contrast to my unbiased sample), that people realize that the uncanny valley is part of their emotions. With further technology advances over the next 10 to 20 years, we expect that technology will be able to render CG humans, which are no longer distinguishable from real humans. However, I would assume that the uncanny valley, Mori's term and its notion will further exist as our intuition will always tell us when something feels not entirely right. While writing this thesis, I observed that people discover the uncanny valley in many aspects of their life. It helps people to express themselves a bit better, and scientific research can pick up these aspects to learn more about humans and human emotions.

[1] `https://longreads.com/2017/11/22/the-uncanny-valley-of-online-dating/`

Apropos emotions, some researchers such as Karl MacDorman as well as Angela Tinwell from the field asked me the third question about my PhD. They asked me:

“Are you enjoying it?”

Today, however, I know that this was the most important question of all during this thesis. The answer lies within the fascination for a counter-intuitive and opportunistic phenomenon that evades human expectations. Before I started my PhD, I worked as a computer graphics and 3D artist on games, films, and documentaries in visual effects (VFX) companies. It was relatively simple and easy for us to render realistic objects or scenes. However, I have always been told that real people are the pinnacle of 3D art, and even when you get millions of dollars, you still need years to render a CG human which is not distinguishable from a real human anymore. During this time, I was not aware of the uncanny valley. It was not even described for CG characters and similarly to the people I am today telling about the uncanny valley, I was not able to describe the actual problem. I thought that there was a fundamental aspect of humans hidden in a mind trick, which I first considered as a bad thing.

During my PhD, I learned much more about the uncanny valley and that humans are hard to create because every individual is as he or she is: perfect and real. Every human is impeccable, immaculate, and unique. And even in times of the Internet, copy-and-paste, machine learning, and artificial intelligence, the human appearance cannot easily be copied or created out-of-the-box because there is something that is not easily replicable. Of course, everyone with a computer is able to manipulate people on photos and make them somehow more appealing (or not). Nevertheless, even highly skilled artists are not able to create new identities or new emotions on humans either in 2D nor 3D. Even when you use immense computing power to render people in 3D, we are currently not able to render complete humans perfectly. For me, the uncanny valley seems to be an opportunistic trend against the total digitization of humans and their world. It is a kind of artifact that distinguishes humans from anything else and it makes every one of us a little bit more special. And yes, learning that was a good feeling, I enjoyed it.

Bibliography

[1] A. Y. Al-Hassan. “A Compendium on the Theory and Practice of the Mechanical Arts.” In: *Journal for the History of Arabic Science Aleppo* 1.1 (1977), pp. 47–64 (cit. on p. 54).

[2] E. L. Altschuler. “Play with online virtual pets as a method to improve mirror neuron and real world functioning in autistic children.” In: *Medical Hypotheses* 70.4 (2008), pp. 748 –749. ISSN: 0306-9877. DOI: `10.1016/j.mehy.2007.07.030` (cit. on p. 153).

[3] W. S. Anderson. *Ovid’s Metamorphoses*. Vol. 2. American Philological Association Series Books 6-10. University of Oklahoma Press, 1972, p. 540. ISBN: 978-0-806-11456-9 (cit. on p. 49).

[4] M. Andrews, B. Chapman. *Brave!* Walt Disney Pictures. 2012 (cit. on p. 31).

[5] K. Apostolakis, P. Daras. “RAAT - The reverie avatar authoring tool.” In: *Proceedings of the 18th International Conference on Digital Signal Processing*. DSP ’13. IEEE. 2013, pp. 1–6. DOI: `10.1109/ICDSP.2013.6622788` (cit. on p. 119).

[6] S. T. Aquinas. *Summa Theologica, Volume 1*. Teil 1. Cosimo Classics, 2013. ISBN: 978-1-602-06553-6 (cit. on p. 55).

[7] F. Argelaguet, L. Hoyet, M. Trico, A. Lécuyer. “The role of interaction in virtual embodiment: Effects of the virtual hand representation.” In: *IEEE Virtual Reality*. IEEE VR ’16. 2016, pp. 3–10. DOI: `10.1109/VR.2016.7504682` (cit. on pp. 91, 190, 206).

[8] J. Axelsson, T. Sundelin, M. J. Olsson, K. Sorjonen, C. Axelsson, J. Lasselin, M. Lekander. “Identification of acutely sick people and facial cues of sickness.” In: *Proceedings of the Royal Society B: Biological Sciences* 285.1870 (2018), p. 20172430. ISSN: 0962-8452. DOI: `10.1098/rspb.2017.2430` (cit. on p. 83).

[9] M. Azmandian, M. Hancock, H. Benko, E. Ofek, A. D. Wilson. “Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences.” In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. Santa Clara, CA, USA: ACM Press, 2016. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858226 (cit. on p. 189).

[10] P. Bader, S. Schneegass, N. Henze, V. Schwind, K. Wolf. “A mobile see-through 3D display with front- and back-touch.” In: *Proceedings of the 8th Nordic Conference on Human-Computer Interaction*. NordiCHI. New York, NY, USA: ACM Press, 2014, pp. 1063–1066. ISBN: 978-1-450-32542-4. DOI: 10.1145/2639189.2670276 (cit. on p. 42).

[11] P. Bader, V. Schwind, N. Henze, S. Schneegass, N. Broy, A. Schmidt. “Design and evaluation of a layered handheld 3d display with touch-sensitive front and back.” In: *Proceedings of the 8th Nordic Conference on Human-Computer Interaction*. NordiCHI. New York, NY, USA: ACM Press, 2014, pp. 315–318. ISBN: 978-1-450-32542-4. DOI: 10.1145/2639189.2639257 (cit. on p. 42).

[12] P. Bader, V. Schwind, N. Pohl, N. Henze, K. Wolf, S. Schneegass, A. Schmidt. “Self-Actuated Displays for Vertical Surfaces.” In: *Interact*. 2015, pp. 282–299. DOI: 10.1007/978-3-319-22723-8_23 (cit. on p. 42).

[13] B. Barnes. *Many Culprits in Fall of a Family Film*. New York Times. New York, NY, USA, 2011 (cit. on p. 67).

[14] S. Baron, H. Widman, L. Lundström. *Real Humans (Äkta människor)*. Matador Film AB/SVT/DT/YLE. 2012 (cit. on p. 66).

[15] C. Bartneck, J. Reichenbach, A. V. Breemen. “In your face, robot! The influence of a character’s embodiment on how users perceive its emotional expressions.” In: *In design and emotion 2004 conference* 2004 (2004), pp. 1–19 (cit. on p. 154).

[16] C. Bartneck, T. Kanda, H. Ishiguro, N. Hagita. “Is the uncanny valley an uncanny cliff?” In: *IEEE International Workshop on Robot and Human Interactive Communication*. IEEE, 2007, pp. 368–373. ISBN: 1-424-41634-5. DOI: 10.1109/ROMAN.2007.4415111 (cit. on pp. 30, 73, 74).

[17] C. Bartneck, D. Kulić, E. Croft, S. Zoghbi. “Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots.” In: *International Journal of Social Robotics* 1.1 (2008), pp. 71–81. ISSN: 1875-4805. DOI: 10.1007/s12369-008-0001-3 (cit. on pp. 74, 155, 156, 158).

[18] C. Bartneck, T. Kanda, H. Ishiguro, H. Norihiro. “My robotic doppelgänger - a critical look at the Uncanny Valley.” In: *The 18th IEEE International Symposium on Robot and Human Interactive Communication*. RO-MAN ’09. 2009, pp. 269–276. DOI: 10.1109/ROMAN.2009.5326351 (cit. on p. 73).

[19] L. F. Baum. *The Wonderful Wizard of Oz*. Oxford University Press, 2008 (cit. on p. 65).

[20] J. M. Beale, F. C. Keil. “Categorical effects in the perception of faces.” In: *Cognition* 57.3 (1995), pp. 217 –239. ISSN: 0010-0277. DOI: https://doi.org/10.1016/0010-0277(95)00669-X (cit. on p. 76).

[21] F. Biocca, M. R. Levy. *Communication in the Age of Virtual Reality*. Routledge Communication Series. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 2013, p. 416. ISBN: 978-1-135-69357-2 (cit. on p. 186).

[22] O. Blanke, T. Metzinger. “Full-body illusions and minimal phenomenal selfhood.” In: *Trends in Cognitive Sciences* 13.1 (2009), pp. 7–13. ISSN: 1364-6613. DOI: 10.1016/j.tics.2008.10.003 (cit. on p. 188).

[23] M. K. Bomers, M. A. van Agtmael, H. Luik, M. C. van Veen, C. M.J. E. Vandenbroucke-Grauls, Y. M. Smulders. “Using a dog’s superior olfactory sensitivity to identify Clostridium difficile in stools and patients: proof of principle study.” In: *BMJ* 345.dec13 8 (2012), e7396–e7396. ISSN: 1756-1833. DOI: 10.1136/bmj.e7396 (cit. on p. 83).

[24] M. Botvinick, J. Cohen. “Rubber hands’ feel’touch that eyes see.” In: *Nature* 391.6669 (1998), p. 756. DOI: 10.1016/j.concog.2007.01.001 (cit. on pp. 188, 190, 220, 221).

[25] A. Bouc. *Men Who Dressup as Rubber Dolls Star on Secrets of the Living Dolls*. guardiantv.com. 2014 (cit. on p. 50).

[26] D. Bovet, J. Vauclair. “Picture Recognition in Animals and Humans.” In: *Behavioural Brain Research* 109.2 (2000), pp. 143–165. ISSN: 0166-4328. DOI: 10.1016/S0166-4328(00)00146-7 (cit. on p. 113).

[27] Box Office Mojo. *"Box Office: Brave"*. boxofficemojo.com. 2012 (cit. on p. 31).

[28] Box Office Mojo. *"Box Office: Mars Needs Moms"*. boxofficemojo.com. 2011 (cit. on p. 31).

[29] E. A. Braemer. *Goethes Prometheus und die Grundpositionen des Strum und Drang*. Beiträge zur deutschen Klassik. Arion Verlag, 1959 (cit. on p. 49).

[30] C. L. Breazeal. *Designing Sociable Robots*. MIT Press, 2004, p. 263. ISBN: 978-0-262-52431-5 (cit. on p. 66).

[31] H. Brenton, M. Gillies, D. Ballin, D. Chatting. "The Uncanny Valley: does it exist?" In: *The 19th British HCI Group Annual Conference: Workshop on Human-animated Character Interaction*. 2005 (cit. on pp. 46, 80, 91, 112).

[32] T. J. Burleigh, J. R. Schoenherr. "A reappraisal of the uncanny valley:categorical perception or frequency-based sensitization?" In: *Frontiers in Psychology* 5.January (2015), pp. 1–19. DOI: 10.3389/fpsyg.2014.01488 (cit. on pp. 93, 179–181, 256).

[33] T. J. Burleigh, J. R. Schoenherr, G. L. Lacroix. "Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces." In: *Computers in Human Behavior* 29.3 (2013), pp. 759–771. ISSN: 0747-5632. DOI: 10.1016/j.chb.2012.11.021 (cit. on pp. 74, 76, 77, 92, 144, 179).

[34] T. Burr. *Movie Review - The Smurfs and friends (sort of) take Manhattan*. Boston.com Globe Newspaper Company. 2011 (cit. on pp. 31, 151).

[35] F. Cafaro, L. Lyons, J. Roberts, J. Radinsky. "The uncanny valley of embodied interaction design." In: *Proceedings of the 2014 Conference on Designing Interactive Systems*. DIS '14. New York, NY, USA: ACM Press, 2014, pp. 1075–1078. ISBN: 978-1-450-32902-6. DOI: 10.1145/2598510.2598593 (cit. on p. 259).

[36] J. Cameron. *Terminator 2: Judgment Day*. Lightstorm Entertainment, Columbia Pictures. 1991 (cit. on p. 20).

[37] J. Cameron. *The Terminator*. Orion Pictures, Hemdale Film Corporation, Pacific Western Productions. 1984 (cit. on p. 67).

[38] R. Campbell, O. Pascalis, M. Coleman, S. B. Wallace, P. J. Benson. "Are faces of different species perceived categorically by human observers?" In: *Proceedings of Biological sciences/The Royal Society* 264.1387 (1997), pp. 1429–1434. ISSN: 0962-8452. DOI: 10.1098/rspb.1997.0199 (cit. on p. 256).

[39] D. Chattopadhyay, K. F. MacDorman. "Familiar faces rendered strange: Why inconsistent realism drives characters into the uncanny valley." In: *Journal of Vision* 16.11 (2016), p. 7. ISSN: 1534-7362. DOI: 10.1167/16.11.7 (cit. on pp. 73, 78–80, 83, 84, 92, 93, 96, 113, 145, 238, 254, 255).

[40] M. Cheetham. "The human likeness dimension of the "uncanny valley hypothesis": behavioral and functional MRI findings." In: *Frontiers in Human Neuroscience* 5 (2011), p. 126. ISSN: 1662-5161. DOI: `10.3389/fnhum.2011.00126` (cit. on pp. 67, 76, 77, 93, 150, 179, 180).

[41] M. Cheetham, L. Jancke. "Perceptual and category processing of the Uncanny Valley hypothesis' dimension of human likeness: some methodological issues." In: *Journal of Visualized Experiments: JoVE* 76 (2013), pp. 1–15. ISSN: 1940-087X. DOI: `10.3791/4375` (cit. on pp. 77, 93, 150, 155, 156, 158, 171).

[42] M. Cheetham, P. Suter, L. Jancke. "Perceptual discrimination difficulty and familiarity in the Uncanny Valley: more like a "Happy Valley"." In: *Frontiers in Psychology* 5 (2014). ISSN: 1664-1078. DOI: `10.3389/fpsyg.2014.01219` (cit. on pp. 76, 77, 93).

[43] M. Cheetham, I. Pavlovic, N. Jordan, P. Suter, L. Jancke. "Category processing and the human likeness dimension of the uncanny valley hypothesis: Eye-tracking data." In: *Frontiers in Psychology* 4 (2013), p. 108. ISSN: 1664-1078. DOI: `10.3389/fpsyg.2013.00108` (cit. on pp. 67, 77, 93, 97, 98, 150, 156, 180, 181).

[44] Z.-H. Chen, C.-Y. Chou, Y.-C. Deng, T.-W. Chan. "Active Open Learner Models As Animal Companions: Motivating Children to Learn Through Interacting with My-Pet and Our-Pet." In: *International Journal of Artificial Intelligence* 17.2 (Apr. 2007), pp. 145–167. ISSN: 1560-4292 (cit. on p. 152).

[45] Z.-H. Chen, C. Liao, T.-C. Chien, T.-W. Chan. "Animal companions: Fostering children's effort-making by nurturing virtual pets." In: *British Journal of Educational Technology* 42.1 (2011), pp. 166–180. ISSN: 1467-8535. DOI: `10.1111/j.1467-8535.2009.01003.x` (cit. on pp. 153, 156).

[46] Z.-H. Chen, Y.-C. Deng, C.-Y. Chou, T.-W. Chan. "Motivating Learners by Nurturing Animal Companions: My-Pet and Our-Pet." In: *Proceedings of the 2005 Conference on Artificial Intelligence in Education: Supporting Learning Through Intelligent and Socially Informed Technology*. IOS Press, 2005, pp. 136–143. ISBN: 1-58603-530-4 (cit. on p. 153).

[47] C. Christou, D. Michael. "Aliens versus Humans: Do Avatars Make a Difference in How We Play the Game?" In: *Proceedings of the 6th International Conference on Games and Virtual Worlds for Serious Applications*. VS-GAMES '14. 2014, pp. 1–7. DOI: `10.1109/VS-Games.2014.7012029` (cit. on p. 91).

[48] D. Chung, B. DeBuys, C. Nam. “Influence of Avatar Creation on Attitude, Empathy, Presence, and Para-Social Interaction.” In: *Proceedings of the 12th International Conference on HCI, Interaction Design and Usability*. 2007, pp. 711–720. ISBN: 978-3-540-73104-7. DOI: 10.1007/978-3-540-73105-4 (cit. on p. 117).

[49] A. C. Clarke. *2001: A Space Odyssey*. Hutchinson (UK)/New American Library (US), 1968, p. 221 (US)/p. 224 (UK). ISBN: 978-1-423-33662-4. DOI: 10.2307/1319100 (cit. on p. 67).

[50] P. Clinton. *"Review: 'Polar Express' a creepy ride"*. CNN.com. 2004 (cit. on p. 31).

[51] J. Cole. “Empathy Needs a Face.” In: *Journal of Consciousness Studies* 8.5-7 (2001), pp. 5–7 (cit. on p. 85).

[52] T. Collins. *Can you tell if she's a robot or a human? 'Life-like android' at a Tokyo gaming conference baffles social media*. MailOnline. 2017 (cit. on p. 32).

[53] N. J. Conard. “A female figurine from the basal Aurignacian of Hohle Fels Cave in southwestern Germany.” In: *Nature* 459.7244 (2009), pp. 248–252. ISSN: 0028-0836. DOI: 10.1038/nature07995 (cit. on p. 19).

[54] M. Costantini, P. Haggard. “The rubber hand illusion: Sensitivity and reference frame for body ownership.” In: *Consciousness and Cognition* 16.2 (2007), pp. 229 –240. ISSN: 1053-8100. DOI: 10.1016/j.concog.2007.01.001 (cit. on p. 220).

[55] M. R. Cunningham. “Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty.” In: *Journal of Personality and Social Psychology* 50.5 (1986), pp. 925–935. ISSN: 0022-3514. DOI: 10.1037/0022-3514.50.5.925 (cit. on p. 127).

[56] T. J. D'Zurilla, G. T. Wilson, R. Nelson. “A preliminary study of the effectiveness of graduated prolonged exposure in the treatment of irrational fear.” In: *Behavior Therapy* 4.5 (1973), pp. 672–685. ISSN: 0005-7894. DOI: 10.1016/S0005-7894(73)80159-5 (cit. on p. 155).

[57] C. Darwin. *The descent of man, and selection in relation to sex*. Vol. 2. D. Appleton, 1872 (cit. on p. 86).

[58] F. de Vignemont. “Embodiment, ownership and disownership.” In: *Consciousness and Cognition* 20.1 (2011), pp. 82–93. ISSN: 1053-8100. DOI: 10.1016/j.concog.2010.09.004 (cit. on p. 188).

[59] P. K. Dick. *Do Androids Dream of Electric Sheep?* Orion, 1968, p. 31. ISBN: 0-345-40447-5 (cit. on p. 65).

[60] T. Dingler, R. Rzayev, V. Schwind, N. Henze. “RSVP on the go - Implicit Reading Support on Smart Watches Through Eye Tracking.” In: *Proceedings of the 2016 ACM International Symposium on Wearable Computers*. ISWC ’16. New York, NY, USA: ACM Press, 2016, pp. 116–119. ISBN: 978-1-450-34460-9. DOI: 10.1145/2971763.2971794 (cit. on p. 42).

[61] A. Dinter. *Der Pygmalion-Stoff in der europäischen Literatur*. Studien zum Fortwirken der Antike. Winter Verlag, 1979, p. 171. ISBN: 978-3-533-02775-1 (cit. on p. 50).

[62] E. Ditzian. *’Tron Legacy’ and Jeff Bridges: Behind Film’s Most Challenging Effect*. MTV News. 2010 (cit. on p. 252).

[63] C. Dormann, J. R. Whitson, M. Neuvians. “Once More With Feeling Game Design Patterns for Learning in the Affective Domain.” In: *Games and Culture* (2013). DOI: 10.1177/1555412013496892 (cit. on p. 153).

[64] E. Downs, S. Smith. “Keeping Abreast of Hypersexuality: A Video Game Character Content Analysis.” English. In: *Sex Roles* 62.11-12 (2010), pp. 721–733. ISSN: 0360-0025. DOI: 10.1007/s11199-009-9637-1 (cit. on p. 117).

[65] N. Ducheneaut, M.-H. Wen, N. Yee, G. Wadley. “Body and Mind: A Study of Avatar Personlization in Three Virtual Worlds.” In: *Proceedings of the 2009 CHI Conference on Human Factors in Computing Systems*. CHI ’09. Boston, MA, USA, 2009, p. 1151. ISBN: 978-1-605-58246-7. DOI: 10.1145/1518701.1518877 (cit. on pp. 118, 146, 251).

[66] A. Duchowski. *Eye Tracking Methodology*. London: Springer London, 2007. ISBN: 978-1-84628-608-7. DOI: 10.1007/978-1-84628-609-4 (cit. on p. 96).

[67] B. R. Duffy. “Anthropomorphism and the social robot.” In: *Robotics and Autonomous Systems* 42.3-4 (2003), pp. 177–190. ISSN: 0921-8890. DOI: 10.1016/S0921-8890(02)00374-3 (cit. on p. 90).

[68] A. Duralde. *‘The Smurfs 2’ Review: What’s With All the Clunky Adult Irony? These Are Just Smurfs!* The Wrap. 2013 (cit. on p. 151).

[69] D. Dutton. “Aesthetics and evolutionary psychology.” In: *The Oxford handbook for aesthetics* (2003), pp. 693–705 (cit. on p. 86).

[70] G. Edwards. *Rogue One: A Star Wars Story*. Lucasfilm Ltd./Walt Disney Pictures. 2016 (cit. on pp. 31, 259).

[71] R. Eisinga, M. T. Grotenhuis, B. Pelzer. “The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown?” In: *International Journal of Public Health* 58.4 (2013), pp. 637–642. ISSN: 1661-8556. DOI: 10.1007/s00038-012-0416-3 (cit. on p. 163).

[72] N. J. Emery. “The eyes have it: The neuroethology, function and evolution of social gaze.” In: *Neuroscience & Biobehavioral Reviews* 24.6 (2000), pp. 581–604. ISSN: 0149-7634. DOI: 10.1016/S0149-7634(00)00025-7 (cit. on p. 113).

[73] M. J. Farah, K. D. Wilson, M Drain, J. N. Tanaka. “What is "special" about face perception?” In: *Psychological Review* 105.3 (1998), pp. 482–498. ISSN: 0033-295X. DOI: 10.1037/0033-295X.105.3.482 (cit. on p. 97).

[74] M. J. Farah. “Is face recognition ’special’? Evidence from neuropsychology.” In: 76.1-2 (1996), pp. 181–189. ISSN: 0166-4328. DOI: 10.1016/0166-4328(95)00198-0 (cit. on p. 97).

[75] M. J. Farah, K. D. Wilson, H. Maxwell Drain, J. R. Tanaka. “The inverted face inversion effect in prosopagnosia: Evidence for mandatory, face-specific perceptual mechanisms.” In: *Vision Research* 35.14 (1995), pp. 2089–2093. ISSN: 0042-6989. DOI: 10.1016/0042-6989(94)00273-O (cit. on p. 97).

[76] J. Favreau. *The Jungle Book*. Walt Disney Pictures. 2016 (cit. on p. 172).

[77] A. M. Feit, D. Weir, A. Oulasvirta. “How We Type: Movement Strategies and Performance in Everyday Typing.” In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. Santa Clara, CA, USA: ACM, 2016, pp. 4262–4273. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858233 (cit. on pp. 225, 236).

[78] N. H. Feldman, T. L. Griffiths, J. L. Morgan. “The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference.” In: *Psychological Review* 116.4 (2009), pp. 752–782. ISSN: 1939-1471. DOI: 10.1037/a0017196 (cit. on p. 78).

[79] A. Felnhofer, O. D. Kothgassner, L. Beutl, H. Hlavacs, I. Kryspin-exner. “Is Virtual Reality made for Men only? Exploring Gender Differences in the Sense of Presence of Psychology, Working Group Clinical and Health Psychology , University of Vienna.” In: *Proceedings of the International Society on Presence Research, Philadelphia, USA* OCTOBER (2012) (cit. on p. 238).

[80] A. Felnhofer, O. D. Kothgassner, N. Hauk, L. Beutl, H. Hlavacs, I. Kryspin-Exner. "Physical and social presence in collaborative virtual environments: Exploring age and gender differences with respect to empathy." In: *Computers in Human Behavior* 31 (2013), pp. 272–279. DOI: 10.1016/j.chb.2013.10.045 (cit. on pp. 187, 238).

[81] M. J. Fenske, J. E. Raymond, K. Kessler, N. Westoby, S. P. Tipper. "Attentional inhibition has social-emotional consequences for unfamiliar faces." In: *Psychological Science* 16.10 (2005), pp. 753–8. ISSN: 0956-7976. DOI: 10.1111/j.1467-9280.2005.01609.x (cit. on p. 82).

[82] A. E. Ferrey, T. J. Burleigh, M. J. Fenske. "Stimulus-category competition, inhibition, and affective devaluation: a novel account of the uncanny valley." In: *Frontiers in Psychology* 6.March (2015), pp. 1–15. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2015.00249 (cit. on pp. 81, 82, 154, 156, 158, 179–181).

[83] L. M. Flach, V. Dill, R. Hocevar, C. Lykawka, S. R. Musse, M. S. Pinho. "Evaluation of the uncanny valley in CG characters." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 7502 LNAI. 2012, pp. 511–513. ISBN: 978-3-642-33196-1. DOI: 10.1007/978-3-642-33197-8-62 (cit. on p. 97).

[84] T. Fong, I. Nourbakhsh, K. Dautenhahn. "A survey of socially interactive robots." In: *Robotics and Autonomous Systems* 42.3-4 (2003), pp. 143–166. ISSN: 0921-8890. DOI: 10.1016/S0921-8890(02)00372-X (cit. on p. 68).

[85] C. Freedman. *Critical Theory and Science Fiction*. Literary studies: Science fiction. Wesleyan University Press, 2000. ISBN: 978-0-819-56399-6 (cit. on p. 61).

[86] S. Freud. "Das Unheimliche." In: *Imago. Zeitschrift für Anwendung der Psychoanalyse auf die Geisteswissenschaften* 1919 (1919), S. 297–324 (cit. on pp. 33, 34, 60, 61, 65).

[87] A. Frischen, A. E. Ferrey, D. H. R. Burt, M. Pistchik, M. J. Fenske. "The affective consequences of cognitive inhibition: Devaluation or neutralization?" In: *Journal of Experimental Psychology: Human Perception and Performance* 38.1 (2012), pp. 169–179. ISSN: 1939-1277. DOI: 10.1037/a0025981 (cit. on p. 81).

[88] J. Frommel, K. Rogers, J. Brich, D. Besserer, L. Bradatsch, I. Ortinau, R. Schabenberger, V. Riemer, C. Schrader, M. Weber. "Integrated Questionnaires: Maintaining Presence in Game Environments for Self-Reported Data Acquisition." In: *Pro-*

ceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play. CHI PLAY '15. London, United Kingdom: ACM, 2015, pp. 359–368. ISBN: 978-1-4503-3466-2. DOI: 10.1145/2793107.2793130 (cit. on p. 194).

[89] J. M. B. Fugate. "Categorical Perception for Emotional Faces." In: *Emotion Review* 5.1 (2013), pp. 84–89. DOI: 10.1177/1754073912451350 (cit. on p. 76).

[90] N. Gaißert, S. Waterkamp, R. W. Fleming, I. Bülthoff. "Haptic Categorical Perception of Shape." In: *PLOS ONE* 7.8 (Aug. 2012), pp. 1–7. DOI: 10.1371/journal.pone.0043062 (cit. on p. 76).

[91] V. Gallese, M. N. Eagle, P. Migone. "Intentional attunement: mirror neurons and the neural underpinnings of interpersonal relations." In: *Journal of the American Psychoanalytic Association* 55.1 (2007), pp. 131–176. ISSN: 0003-0651. DOI: 10.1177/00030651070550010601 (cit. on p. 85).

[92] S. W. Gangestad, D. M. Buss. "Pathogen prevalence and human mate preferences." In: *Ethology and Sociobiology* 14 (1993), pp. 89–96. ISSN: 0162-3095. DOI: 10.1016/0162-3095(93)90009-7 (cit. on pp. 87, 120).

[93] A. Garcia-Palacios, H. Hoffman, A. Carlin, T. Furness, C. Botella. "Virtual reality in the treatment of spider phobia: a controlled study." In: *Behaviour Research and Therapy* 40.9 (2002), pp. 983–993. ISSN: 0005-7967. DOI: 10.1016/S0005-7967(01)00068-7 (cit. on p. 153).

[94] H. G. Gassen, S Minol. *Die Menschen Macher: Sehnsucht nach Unsterblichkeit*. Erlebnis Wissenschaft. Wiley, 2012. ISBN: 978-3-527-64092-8 (cit. on p. 49).

[95] P. Gendolla. *Anatomien der Puppe: Zur Geschichte des MaschinenMenschen bei Jean Paul, E. T. A. Hoffman, Villiers de L'Isle-Adam und Hans Bellmer*. Vol. 113. Universitätsverlag Winter, 1992, p. 239. ISBN: 978-3-533-04556-4 (cit. on p. 57).

[96] L. Giles. *Taoist teachings from the book of Lieh Tzŭ*. Wisdom of the East series Bd. 6. J. Murray, 1925 (cit. on p. 52).

[97] M. J. Giummarra, N. Georgiou-Karistianis, M. E. R. Nicholls, S. J. Gibson, J. L. Bradshaw. "The Phantom in the Mirror: A Modified Rubber-Hand Illusion in Amputees and Normals." In: *Perception* 39.1 (2010). PMID: 20301851, pp. 103–118. DOI: 10.1068/p6519 (cit. on p. 188).

[98] J. W. V. Goethe. *Prometheus: Dramatisches Fragment*. Library of Alexandria, 1789. ISBN: 978-1-465-51395-3 (cit. on p. 49).

[99] R. L. Goldstone, A. T. Hendrickson. “Categorical perception.” In: *Wiley Interdisciplinary Reviews: Cognitive Science* 1.1 (2010), pp. 69–78. ISSN: 1939-5086. DOI: 10.1002/wcs.26 (cit. on p. 76).

[100] R. Gosnell. *The Smurfs I*. Sony Pictures Animation, Columbia Pictures. 2011 (cit. on p. 172).

[101] A. C. Graham. *The book of Lieh-tzu:A Classic of The Tao*. Unesco collection of representative works. Chinese series. J. Murray, 1990, xix, 192 p. ISBN: 978-0-231-07236-6 (cit. on p. 52).

[102] K. Grammer, R. Thornhill. “Human (Homo sapiens) facial attractiveness and sexual selection: the role of symmetry and averageness.” In: *Journal of Comparative Psychology* 108.3 (1994), pp. 233–242. ISSN: 0735-7036. DOI: 10.1037/0735-7036.108.3.233 (cit. on pp. 120, 123).

[103] K. Gray, D. M. Wegner. “Feeling robots and human zombies:Mind perception and the uncanny valley.” In: *Cognition* 125.1 (2012), pp. 125–130. ISSN: 0010-0277. DOI: 10.1016/j.cognition.2012.06.007 (cit. on pp. 80, 150).

[104] J. Greenberg, T. Pyszczynski, S. Solomon. “The Causes and Consequences of a Need for Self-Esteem: A Terror Management Theory.” In: *Public Self and Private Self*. Ed. by R. F. Baumeister. New York, NY: Springer New York, 1986, pp. 189–212. ISBN: 978-1-4613-9564-5. DOI: 10.1007/978-1-4613-9564-5_10 (cit. on p. 84).

[105] O. Griffiths, C. J. Mitchell. “Negative priming reduces affective ratings.” In: *Cognition & Emotion* 22.6 (2008), pp. 1119–1129. ISSN: 0269-9931. DOI: 10.1080/02699930701664930 (cit. on p. 82).

[106] P. Gutierrez. “Identity and the Uncanny: A Coraline Study Guide.” In: *Screen Education* 56 (2009), pp. 84–89 (cit. on p. 151).

[107] D. Hanson. “Exploring the aesthetic range for humanoid robots.” In: *Proceedings of the ICCS/CogSci-2006 long symposium: Toward social mechanisms of android science* (2006), pp. 39–42 (cit. on pp. 50, 87, 88).

[108] D. Hanson, A. Olney, S Prilliman. “Upending the uncanny valley.” In: *Proceedings of the 20th National Conference on Artificial Intelligence* 20.July (2005), pp. 24–31 (cit. on pp. 50, 69, 74, 88).

[109] D. E. Harkness. *Promethean Ambitions: Alchemy and the Quest to Perfect Nature (review)*. Vol. 48. American Politics and Political Economy Series. University of Chicago Press, 2007, pp. 192–193. DOI: 10.1353/tech.2007.0018 (cit. on p. 55).

[110] S. Harnad. "Psychophysical and cognitive aspects of categorical perception: A critical overview." In: *Categorical Perception: The Groundwork of Cognition*. Cambridge University Press, 1987, pp. 1–52 (cit. on p. 76).

[111] S. G. Hart. "NASA-task load index (NASA-TLX); 20 years later." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 50. 9. Sage Publications. 2006, pp. 904–908. DOI: 10.1177/154193120605000909 (cit. on p. 227).

[112] J. V. Haxby, E. a. Hoffman, M. Gobbini. "Human neural systems for face recognition and social communication." In: *Biological Psychiatry* 51.1 (2002), pp. 59–67. ISSN: 0006-3223. DOI: 10.1016/S0006-3223(01)01330-0 (cit. on p. 97).

[113] H. Heckmann. *Die andere Schöpfung: Geschichte der frühen Automaten in Wirklichkeit und Dichtung*. Umschau, 1982 (cit. on p. 61).

[114] C. Heeter, R. Egidio, P. Mishra, B. Winn, J. Winn. "Alien Games: Do Girls Prefer Games Designed by Girls?" In: *Games and Culture* 4 (2008), pp. 74–100. ISSN: 1555-4120. DOI: 10.1177/1555412008325481 (cit. on p. 118).

[115] C. S. Henshilwood, F. d'Errico, K. L. van Niekerk, Y. Coquinot, Z. Jacobs, S.-E. Lauritzen, M. Menu, R. García-Moreno. "A 100,000-Year-Old Ochre-Processing Workshop at Blombos Cave, South Africa." In: *Science* 334.6053 (2011), pp. 219–222. ISSN: 0036-8075. DOI: 10.1126/science.1211535 (cit. on p. 18).

[116] N. Henze, S. Mayer, H. V. Le, V. Schwind. "Improving software-reduced touchscreen latency." In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '17. New York, NY, USA: ACM Press, 2017, pp. 1–8. ISBN: 9781450350754. DOI: 10.1145/3098279.3122150 (cit. on p. 42).

[117] Hesiod, T. A. Sinclair. *Hesiod Works and Days (Greek Texts and Commentaries)*. 1979, p. 96. ISBN: 978-0-405-11446-5 (cit. on p. 49).

[118] D. R. Hill. "Mechanical engineering in the medieval near east." In: *Scientific American* 264.5 (1991), pp. 100–105 (cit. on p. 19).

[119] D. Hill. *A History of Engineering in Classical and Medieval Times*. Vol. 1996. Routledge, 1996, p. 263. ISBN: 0-415-15291-7 (cit. on p. 54).

[120] C. C. Ho, K. F. MacDorman. "Measuring the Uncanny Valley Effect: Refinements to Indices for Perceived Humanness, Attractiveness, and Eeriness." In: *International Journal of Social Robotics* 9.1 (2017), pp. 129–139. ISSN: 1875-4805. DOI: 10.1007/s12369-016-0380-9 (cit. on p. 74).

[121] C. C. Ho, K. F. MacDorman. "Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices." In: *Computers in Human Behavior* 26.6 (2010), pp. 1508–1518. ISSN: 0747-5632. DOI: 10.1016/j.chb.2010.05.015 (cit. on pp. 74, 155, 158).

[122] C. C. Ho, K. F. MacDorman, Z. Pramono. "Human emotion and the uncanny valley: a GLM, MDS, and Isomap analysis of robot video ratings." In: *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*. ACM/IEEE HRI '08. 2008, pp. 169–176. ISBN: 978-1-605-58017-3. DOI: 10.1145/1349822.1349845 (cit. on pp. 74, 97).

[123] E. Hoffman. *Der Sandmann*. Reclam, 2008. ISBN: 978-3-15-000230-8 (cit. on pp. 33, 60).

[124] N. Holt. *Guys and Dolls*. BBC. 2007 (cit. on p. 50).

[125] L. Hoyet, F. Argelaguet, A. Lécuyer. ""Wow! I Have Six Fingers!": Would You Accept Structural Changes of Your Hand in VR?" In: *Frontiers in Robotics and AI* 3 (2016), p. 27. ISSN: 2296-9144. DOI: 10.3389/frobt.2016.00027 (cit. on pp. 188, 190, 221).

[126] J. Hsu. *Robotics' Uncanny Valley Gets New Translation*. Live Science. 2012 (cit. on p. 30).

[127] Y. Hswen, V. Murti, A. A. Vormawor, R. Bhattacharjee, J. A. Naslund. "Virtual Avatars, Gaming, and Social Media: Designing a Mobile Health App To Help Children Choose Healthier Food Options." In: *Journal of Mobile Technology in Medicine* 2.2 (2013), pp. 8–14. ISSN: 1839-7808. DOI: 10.7309/jmtm.2.2.3 (cit. on pp. 152, 156).

[128] G. Hulfish. *The uncanny valley effect is strong with 'L.A. Noire' in 4K Ultra HD*. Digital Trends. 2016 (cit. on p. 32).

[129] O. Hurard. *Lars Lundström, Autor der Serie "Real Humans": „Ein Spiegel unserer eigenen Existenz"*. arte.tv. 2013 (cit. on p. 66).

[130] Z. Hussain, M. D. Griffiths. "Gender swapping and socializing in cyberspace: An exploratory study." In: *CyberPsychology & Behavior* 11.1 (2008), pp. 47–53 (cit. on p. 118).

[131] J. P. Huston, M. Nadal, F. Mora, L. F. Agnati, C. J. C. Conde, eds. *Art, Aesthetics, and the Brain*. Oxford University Press, 2015. ISBN: 978-0-199-67000-0. DOI: 10.1093/acprof:oso/9780199670000.001.0001 (cit. on p. 86).

[132] M. Hutson. *The 7 Laws of Magical Thinking: How Irrational Beliefs Keep Us Happy, Healthy, and Sane*. Penguin, 2012, p. 304. ISBN: 978-1-101-55832-4 (cit. on p. 24).

[133] J. Hyde, E. J. Carter, S. Kiesler, J. K. Hodgins. "Perceptual effects of damped and exaggerated facial motion in animated characters." In: FG '13. 2013. ISBN: 978-1-467-35545-2. DOI: 10.1109/FG.2013.6553775 (cit. on p. 182).

[134] R. J. Itier, C. Villate, J. D. Ryan. "Eyes always attract attention but gaze orienting is task-dependent: Evidence from eye movement monitoring." In: *Neuropsychologia* 45.5 (2007), pp. 1019–1028. ISSN: 0028-3932. DOI: 10.1016/j.neuropsychologia.2006.09.004 (cit. on p. 97).

[135] T. Jacobsen. "Beauty and the brain: culture, history and individual differences in aesthetic appreciation." In: *Journal of Anatomy* 216.2 (2010), pp. 184–191. ISSN: 0021-8782. DOI: 10.1111/j.1469-7580.2009.01164.x (cit. on p. 86).

[136] S. Jäger. *Erfolgreiches Charakterdesign für Computer- und Videospiele*. Wiesbaden: Springer Fachmedien Wiesbaden, 2013. ISBN: 978-3-658-02625-7. DOI: 10.1007/978-3-658-02626-4 (cit. on pp. 139, 187).

[137] S. W. Janik, a. R. Wellens, M. L. Goldberg, L. F. Dell'Osso. "Eyes as the center of focus in the visual examination of human faces." In: *Perceptual and Motor Skills* 47 (1978), pp. 857–858. ISSN: 0031-5125. DOI: 10.2466/pms.1978.47.3.857 (cit. on p. 97).

[138] U. P. Jauch. *Jenseits der Maschine: Philosophie, Ironie und Ästhetik bei Julien Offray de La Mettrie (1709-1751)*. C. Hanser Verlag, 1998, p. 598. ISBN: 978-3-446-19485-4 (cit. on p. 59).

[139] E. Jentsch. "Psychiatrisch-Neurologische Wochenschrift." In: *Psychiatrisch-Neurologische Wochenschrift* 22 (1906), pp. 195–205 (cit. on pp. 33, 34).

[140] S. John. "Un/realistically embodied: The gendered conceptions of realistic game design." In: *Online Proceedings of 2006 AVI Workshop „Gender and Interaction"*. 2006 (cit. on pp. 117, 118).

[141] L. Jordan. "Pars Secunda Philosophiae, seu Metaphisica." In: *Archiv für Geschichte der Philosophie* 23 (1-4 1910). ISSN: 0003-9101. DOI: 10.1515/agph.1910.23.3.338 (cit. on p. 59).

[142] M. A. Just, P. A. Carpenter. “Eye fixations and cognitive processes.” In: *Cognitive Psychology* 8.4 (1976), pp. 441–480 (cit. on p. 96).

[143] F. Kaba. “Hyper-Realistic Characters and the Existence of the Uncanny Valley in Animation Films.” In: *International Review of Social Sciences and Humanities* 4.2 (2013), pp. 188–195 (cit. on p. 96).

[144] J. Kätsyri. “"The Polar Express" is Bipolar: Critical Film Reviews Influence Uncanny Valley Phenomenon in Animation Films.” In: *Proceedings of 39th Annual Meeting of the Cognitive Science Society*. 2017 (cit. on p. 31).

[145] J. Kätsyri, K. Förger, M. Mäkäräinen, T. Takala. “A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness.” In: *Frontiers in Psychology* 6.MAR (2015), p. 390. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2015.00390 (cit. on pp. 30, 37, 72–77, 80–82, 84, 88, 89, 93, 144, 150, 155, 156, 170, 180).

[146] T. Kawabe, K. Sasaki, K. Ihaya, Y. Yamada. “When categorization-based stranger avoidance explains the uncanny valley: A comment on MacDorman and Chattopadhyay (2016).” In: *Cognition* (2016). ISSN: 1873-7838. DOI: 10.1016/j.cognition.2016.09.001 (cit. on pp. 82, 154, 156, 179, 181).

[147] D. J. Kelly, P. C. Quinn, A. M. Slater, K. Lee, L. Ge, O. Pascalis. “The Other-Race Effect Develops During Infancy.” In: *Psychological Science* 18.12 (2007), pp. 1084–1089. ISSN: 0956-7976. DOI: 10.1111/j.1467-9280.2007.02029.x (cit. on p. 78).

[148] C. Kidd, W. Taggart, S. Turkle. “A sociable robot to encourage social interaction among the elderly.” In: *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*. IEEE ICRA ’06. IEEE, 2006, pp. 3972–3976. ISBN: 0-7803-9505-0. DOI: 10.1109/ROBOT.2006.1642311 (cit. on p. 151).

[149] K. Kilteni, R. Groten, M. Slater. “The sense of embodiment in virtual reality.” In: *Presence: Teleoperators and Virtual Environments* 21.4 (2012), pp. 373–387 (cit. on p. 186).

[150] P. Knierim, V. Schwind, A. Feit, F. Nieuwenhuizen, N. Henze. “Physical Keyboards in Virtual Reality: Analysis of Typing Performance and Effects of Avatar Hands.” In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. Montréal, Canada: ACM, 2018. ISBN: 978-1-4503-5620-6/18/04. DOI: 10.1145/3173574.3173919 (cit. on pp. 40, 186).

[151] E. Kohler, C. Keysers, M. A. Umiltà, L. Fogassi, V. Gallese, G. Rizzolatti. “Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons.” In: *Science* 297.5582 (2002), pp. 846–848. ISSN: 0036-8075. DOI: 10.1126/science.1070311 (cit. on p. 85).

[152] V. Koreis. *Čapek’s R.U.R.* 2008 (cit. on p. 63).

[153] E. Kormann. “Künstliche Menschen oder der moderne Prometheus. Der Schrecken der Autonomie.” In: *Amsterdamer Beiträge zur neueren Germanistik* 59.1 (2006) (cit. on p. 62).

[154] O. Korn, M. Blatz, A. Rees, J. Schaal, V. Schwind, D. Görlich. “Procedural Content Generation for Game Props? A Study on the Effects on User Experience.” In: *Computers in Entertainment* 15.2 (2017), pp. 1–15. ISSN: 1544-3574. DOI: 10.1145/2974026 (cit. on p. 42).

[155] S. Krach, F. Hegel, B. Wrede, G. Sagerer, F. Binkofski, T. Kircher. “Can Machines Think? Interaction and Perspective Taking with Robots Investigated via fMRI.” In: *PLOS ONE* 3.7 (July 2008), pp. 1–11. DOI: 10.1371/journal.pone.0002597 (cit. on p. 85).

[156] S. Kubrick. *2001: A Space Odyssey*. 1968. ISBN: 978-1-423-33662-4. DOI: 10.2307/1319100 (cit. on p. 67).

[157] K. Lagrandeur. “Do medieval and renaissance androids presage the posthuman?” In: *CLCWeb - Comparative Literature and Culture* 12.3 (2010). ISSN: 1481-4374. DOI: 10.7771/1481-4374.1553 (cit. on p. 55).

[158] F. Lang. *Metropolis*. UFA Film. 1927 (cit. on pp. 63, 64).

[159] J. H. Langlois, L. A. Roggman. “Attractive Faces Are Only Average.” In: *Psychological Science* 1.2 (1990), pp. 115–121. DOI: 10.1111/j.1467-9280.1990.tb00079.x (cit. on pp. 86, 93, 120).

[160] J. H. Langlois, L. A. Roggman, L. Musselman. “What is average and what is not average about attractive faces?” In: *Psychological Science* 5.4 (1994), pp. 214–220 (cit. on pp. 123, 136).

[161] S. R. H. Langten, R. J. Watt, V. Bruce. “Do the eyes have it? Cues to the direction of social attention.” In: *Trends in Cognitive Sciences* 4.2 (2000), pp. 50–59. ISSN: 1364-6613. DOI: 10.1016/S1364-6613(99)01436-9 (cit. on p. 97).

[162] A. Larsson, C. Nerén. “Gender Aspects of Computer Avatars.” In: *SICS Research Report* (2005) (cit. on p. 117).

[163] K. Lawler. *How the 'Rogue One' ending went wrong*. USA Today. 2016 (cit. on pp. 32, 252).

[164] S. Lay, N. Brace, G. Pike, F. Pollick. "Circling Around the Uncanny Valley: Design Principles for Research Into the Relation Between Human Likeness and Eeriness." In: *i-Perception* 7.6 (2016). ISSN: 2041-6695. DOI: 10.1177/2041669516681309 (cit. on pp. 30, 72).

[165] S. Lay. *Meet China's freakishly realistic droids: Creepy machines go on display at the World Robot conference*. The Guardian. 2015 (cit. on p. 32).

[166] H. V. Le, V. Schwind, P. Göttlich, N. Henze. "PredicTouch: A System to Reduce Touchscreen Latency using Neural Networks and Inertial Measurement Units." In: *Proceedings of the Interactive Surfaces and Spaces*. New York, NY, USA: ACM Press, 2017, pp. 230–239. ISBN: 978-1-450-34691-7. DOI: 10.1145/3132272.3134138 (cit. on p. 42).

[167] C. Lecouteux. *Histoire des Vampires, Autopsie d'un mythe*. Editions Imago 1999, 1999. ISBN: 978-2-911416-29-3 (cit. on p. 55).

[168] S. L. Lee, I. Y. M. Lau, S. Kiesler, C. Y. Chiu. "Human mental models of humanoid robots." In: *IEEE International Conference on Robotics and Automation*. Vol. 2005. IEEE, 2005, pp. 2767–2772. ISBN: 0-780-38914-X. DOI: 10.1109/ROBOT.2005.1570532 (cit. on p. 68).

[169] D. A. Leopold, G. Rhodes. "A comparative view of face perception." In: *Journal of Comparative Psychology* 124.3 (2010), pp. 233–251. ISSN: 0735-7036. DOI: 10.1037/a0019460 (cit. on p. 113).

[170] D. J. Lewkowicz, A. A. Ghazanfar. "The development of the uncanny valley in infants." In: *Developmental Psychobiology* 54.2 (2012), pp. 124–132. ISSN: 0012-1630. DOI: 10.1002/dev.20583 (cit. on p. 78).

[171] A. M. Liberman, K. S. Harris, H. S. Hoffman, B. C. Griffith. "The discrimination of speech sounds within and across phoneme boundaries." In: *Journal of Experimental Psychology* 54.4 (1957), pp. 358–368. DOI: 10.1037/h0044417 (cit. on p. 76).

[172] Lieh-Tzu, R. Wilhelm. *Das Wahre Buch Vom Quellenden Urgrund*. Stuttgart: Zenodot Verlagsgesellscha, 1980. ISBN: 978-3-843-06565-8 (cit. on p. 52).

[173] J.-W. Lin, P.-H. Han, J.-Y. Lee, Y.-S. Chen, T.-W. Chang, K.-W. Chen, Y.-P. Hung. "Visualizing the Keyboard in Virtual Reality for Enhancing Immersive Experience." In: *ACM SIGGRAPH 2017 Posters*. SIGGRAPH '17. Los Angeles, CA, USA: ACM, 2017, 35:1–35:2. ISBN: 978-1-4503-5015-0. DOI: 10.1145/3102163.3102175 (cit. on p. 235).

[174] L. Lin, S. Jörg. "Need a Hand?: How Appearance Affects the Virtual Hand Illusion." In: *Proceedings of the ACM Symposium on Applied Perception*. SAP '16. Anaheim, CA, USA: ACM, 2016, pp. 69–76. ISBN: 978-1-4503-4383-1. DOI: 10.1145/2931002.2931006 (cit. on pp. 91, 190).

[175] A. Liptak. *The first trailer for Robert Rodriguez's Alita: Battle Angel falls squarely into the uncanny valley*. The Verge. 2017 (cit. on p. 249).

[176] L. Lischke, V. Schwind, K. Friedrich, N. Henze, A. Schmidt. "MAGIC-Pointing on Large High-Resolution Displays." In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '16. ACM Press, 2016. ISBN: 978-1-450-34082-3. DOI: 10.1145/2851581.2892479 (cit. on p. 42).

[177] M. Lombard, T. Ditton. "At the Heart of It All: The Concept of Presence." In: *Journal of Computer-Mediated Communication* 3.2 (1997). ISSN: 1083-6101. DOI: 10.1111/j.1083-6101.1997.tb00072.x (cit. on p. 187).

[178] M. R. Longo, F. Schüür, M. P. Kammers, M. Tsakiris, P. Haggard. "What is embodiment? A psychometric approach." In: *Cognition* 107.3 (2008), pp. 978–998. ISSN: 0010-0277. DOI: 10.1016/j.cognition.2007.12.004 (cit. on p. 188).

[179] C. E. Looser, T. Wheatley. "The tipping point of animacy. How, when, and where we perceive life in a face." In: *Psychological Science: A Journal of the American Psychological Society/APS* 21.12 (2010), pp. 1854–1862. ISSN: 0956-7976. DOI: 10.1177/0956797610388044 (cit. on pp. 76, 77, 97, 150).

[180] H. J. Lu, L. Chang. "Automatic attention towards face or body as a function of mating motivation." In: *Evolutionary psychology:an international journal of evolutionary approaches to psychology and behavior* 10.1 (2012), pp. 120–135. ISSN: 1474-7049 (cit. on p. 87).

[181] J.-L. Lugrin, J. Latt, M. E. Latoschik. "Anthropomorphism and Illusion of Virtual Body Ownership." In: *Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual*

Environments. ICAT - EGVE '15. Kyoto, Japan: Eurographics Association, 2015, pp. 1–8. ISBN: 978-3-905674-84-2. DOI: 10.2312/egve.20151303 (cit. on pp. 91, 186, 204).

[182] L. Lumenick. *"Box Office: 'Mars Needs Moms' A Megaton Bomb"*. New York Post. 2011 (cit. on p. 31).

[183] K. F. MacDorman. "Androids as an experimental apparatus: Why is there an uncanny valley and can we exploit it." In: *CogSci-2005 workshop: toward social mechanisms of android science* 3 (2005), pp. 106–118 (cit. on pp. 55, 84, 156).

[184] K. F. MacDorman. "Mortality salience and the uncanny valley." In: *Proceedings of 2005 5th IEEE-RAS International Conference on Humanoid Robots* (2005), pp. 399–405. DOI: 10.1109/ICHR.2005.1573600 (cit. on pp. 92, 97, 155, 181).

[185] K. F. MacDorman. "Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley." In: *ICCS/CogSci-2006 long symposium: Toward social mechanisms of android science* (2006), pp. 26–29 (cit. on pp. 92, 145, 156).

[186] K. F. MacDorman, D. Chattopadhyay. "Categorization-based stranger avoidance does not explain the uncanny valley effect." In: *Cognition* (2017), pp. 7–10. ISSN: 1873-7838. DOI: 10.1016/j.cognition.2017.01.009 (cit. on pp. 84, 93, 154).

[187] K. F. MacDorman, D. Chattopadhyay. "Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not." In: *Cognition* 146 (2016), pp. 190–205. ISSN: 0010-0277. DOI: 10.1016/j.cognition.2015.09.019 (cit. on pp. 30, 77, 78, 80, 92, 93, 96, 113, 145, 153, 154, 156, 180, 181, 238, 250, 254, 255).

[188] K. F. MacDorman, S. O. Entezari. "Individual differences predict sensitivity to the uncanny valley." In: *Interaction Studies* (2015), pp. 1–47. DOI: 10.1075/is.16.2.01mac (cit. on p. 155).

[189] K. F. MacDorman, H. Ishiguro. "The uncanny advantage of using androids in cognitive and social science research." In: *Interaction Studies* 7.3 (2006), pp. 297–337. ISSN: 1572-0373. DOI: 10.1075/is.7.3.03mac (cit. on pp. 20, 33).

[190] K. F. MacDorman, P. Srinivas, H. Patel. "The uncanny valley does not interfere with level 1 visual perspective taking." In: *Computers in Human Behavior* 29.4 (2013), pp. 1671–1685. ISSN: 0747-5632. DOI: 10.1016/j.chb.2013.01.051 (cit. on pp. 85, 155, 156).

[191] K. F. MacDorman, S. K. Vasudevan, C. C. Ho. "Does Japan really have robot mania? Comparing attitudes by implicit and explicit measures." In: *AI and Society* 23.4 (2009), pp. 485–510. ISSN: 0951-5666. DOI: 10.1007/s00146-008-0181-2 (cit. on pp. 67, 88).

[192] K. F. MacDorman, R. D. Green, C. C. Ho, C. T. Koch. "Too real for comfort? Uncanny responses to computer generated faces." In: *Computers in Human Behavior* 25.3 (2009), pp. 695–710. ISSN: 0747-5632. DOI: 10.1016/j.chb.2008.12.026 (cit. on pp. 46, 67, 73, 80, 81, 83, 85, 88, 92, 93, 97, 139, 150, 156, 157, 169, 172, 177, 180, 181, 255).

[193] I. S. MacKenzie, R. W. Soukoreff. "Phrase Sets for Evaluating Text Entry Techniques." In: *Proceedings of the 2003 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '03. Ft. Lauderdale, Florida, USA: ACM, 2003, pp. 754–755. ISBN: 1-58113-637-4. DOI: 10.1145/765891.765971 (cit. on p. 226).

[194] M. Mäkäräinen, J. Kätsyri, T. Takala. "Exaggerating Facial Expressions: A Way to Intensify Emotion or a Way to the Uncanny Valley?" In: *Cognitive Computation* 6.4 (2014), pp. 708–721. ISSN: 1866-9956. DOI: 10.1007/s12559-014-9273-0 (cit. on pp. 90, 157, 172, 178).

[195] M. B. Mathur, D. B. Reichling. "Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley." In: *Cognition* 146.Supplement C (2016), pp. 22 –32. ISSN: 0010-0277. DOI: https://doi.org/10.1016/j.cognition.2015.09.008 (cit. on pp. 30, 73, 74, 87, 89, 90, 170).

[196] S. M. Maul. *Das Gilgamesch-Epos*. 5th ed. C.H.Beck, 2012, p. 192. ISBN: 978-3-406-64232-6 (cit. on p. 47).

[197] S. Mayer, K. Wolf, S. Schneegass, N. Henze. "Modeling Distant Pointing for Compensating Systematic Displacements." In: *Proceedings of the 2015 CHI Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, 2015, pp. 4165–4168. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702332 (cit. on p. 258).

[198] S. Mayer, V. Schwind, R. Schweigert, N. Henze. "The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montréal, Canada: ACM, 2018. ISBN: 978-1-4503-5620-6/18/04. DOI: 10.1145/3173574.3174227 (cit. on p. 258).

[199] R. McDonnell, M. Breidt. “Face Reality:Investigating the Uncanny Valley for virtual faces.” In: *ACM SIGGRAPH ASIA 2010 Sketches*. ACM, 2010, p. 41. ISBN: 978-1-450-30523-5. DOI: 10.1145/1899950.1899991 (cit. on p. 93).

[200] R. McDonnell, M. Breidt, H. H. Bülthoff. “Render me Real? Investigating the Effect of Render Style on the Perception of Animated Virtual Humans.” In: *ACM Transactions on Graphics* 31.4 (2012), pp. 1–11. ISSN: 0730-0301. DOI: 10.1145/2185520.2185587 (cit. on pp. 74, 93, 182).

[201] D. E. McDowell, J. Ruby. *Secure the Shadow: Death and Photography in America*. Vol. 26. MIT Press, 1997, p. 377. DOI: 10.2307/2654062 (cit. on p. 98).

[202] M. McGill, D. Boland, R. Murray-Smith, S. Brewster. “A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays.” In: *Proceedings of the 2015 CHI Conference on Human Factors in Computing Systems*. CHI ’15. Seoul, Republic of Korea: ACM, 2015, pp. 2143–2152. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702382 (cit. on pp. 188, 235).

[203] F. J. Mercer Moss, R. Baddeley, N. Canagarajah. “Eye Movements to Natural Images as a Function of Sex and Personality.” In: *PLoS ONE* 7.11 (2012), e47870. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0047870 (cit. on p. 105).

[204] J. O.D. L. Mettrie. *L’Homme Machine - Scholar’s Choice Edition*. 2015. ISBN: 978-1-296-22676-3 (cit. on p. 59).

[205] A. Michaels. *Secrets of the Living Dolls, Channel 4, Review*. The Telegraph. 2014 (cit. on p. 50).

[206] M. K. Miller, A. Summers. “A content analysis of the evolution of video game characters as represented in video game magazines.” In: *Journal of Media Psychology* 14.3 (2009) (cit. on pp. 152, 156).

[207] M. Minear, D. C. Park. “A lifespan database of adult facial stimuli.” In: *Behavior research methods, instruments, & computers: A journal of the Psychonomic Society, Inc* 36 (2004), pp. 630–633. ISSN: 0743-3808. DOI: 10.3758/BF03206543 (cit. on p. 123).

[208] C. Misselhorn. “Empathy with inanimate objects and the uncanny valley.” In: *Minds and Machines* 19.3 (2009), pp. 345–359. ISSN: 0924-6495. DOI: 10.1007/s11023-009-9158-2 (cit. on pp. 67, 85, 97).

[209] W. J. Mitchell, K. A. Szerszen, A. S. Lu, P. W. Schermerhorn, M. Scheutz, K. F. MacDorman. “A Mismatch in the Human Realism of Face and Voice Produces an Uncanny Valley.” In: *i-Perception* 2.1 (2011), pp. 10–12. DOI: 10.1068/i0415 (cit. on pp. 80, 150, 155, 156).

[210] R. K. Moore. “A Bayesian explanation of the ’Uncanny Valley’ effect and related psychological phenomena.” In: *Nature* 2 (2012), p. 864. ISSN: 2045-2322. DOI: 10.1038/srep00864 (cit. on pp. 30, 78–80, 82, 113, 145, 150, 238, 248, 255).

[211] M. Mori. *The Buddha in the Robot*. Ed. by C. Terry, R. Friedrich. Kosei Publishing Company, 1989, p. 192. ISBN: 978-4-333-01002-8 (cit. on p. 68).

[212] M. Mori. “The Uncanny Valley.” In: *Energy* 7.4 (1970), pp. 33–35 (cit. on pp. 25, 26, 46).

[213] M. Mori, K. F. MacDorman, N. Kageki. “The Uncanny Valley.” In: *IEEE Robotics and Automation Magazine* 19.2 (2012), pp. 98–100. ISSN: 1070-9932. DOI: 10.1109/MRA.2012.2192811 (cit. on pp. 25, 26, 29, 72, 75, 80, 83, 87–91, 93, 145, 154, 156, 169, 170, 179, 182, 238).

[214] G. Müller. *Die Utopie in der deutschen Literatur*. J.B. Metzlersche Verlagsbuchhandlung, 1989. ISBN: 978-3-476-00661-5 (cit. on p. 61).

[215] C. S. Murray. “Horror as Dr X builds his creatur.” In: *The Telegraph* (2002) (cit. on p. 61).

[216] C. D. Murray, S. Pettifer, T. Howard, E. L. Patchick, F. Caillette, J. Kulkarni, C. Bamford. “The treatment of phantom limb pain using immersive virtual reality: Three case studies.” In: *Disability and Rehabilitation* 29.18 (2007). PMID: 17729094, pp. 1465–1469. DOI: 10.1080/09638280601107385 (cit. on pp. 188, 221).

[217] M. Mustafa, S. Guthe, J.-P. Tauscher, M. Goesele, M. Magnor. “How Human Am I?: EEG-based Evaluation of Virtual Characters.” In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17. New York, NY, USA: ACM Press, 2017, pp. 5098–5108. ISBN: 978-1-450-34655-9. DOI: 10.1145/3025453.3026043 (cit. on p. 89).

[218] G. Nadarajan. “Islamic Automation:A Reading of al-Jazari’s The Book of Knowledge of Ingenious Mechanical Devices (1206).” In: *Foundation for Science Technology and Civilisation* 1206.803 (2007), pp. 1–16 (cit. on p. 54).

[219] J. Needham, L. Wang. *Science and Civilisation in China*. Vol. 2. History of Scientific Thought. Cambridge University Press, 1956. ISBN: 978-0-521-05800-1 (cit. on p. 52).

[220] A. Newitz. *Is the "uncanny valley" a myth?* io9/Gizmodo. 2013 (cit. on p. 46).

[221] K. L. Nowak, C. Rauh. "The Influence of the Avatar on Online Perceptions of Anthropomorphism, Androgyny, Credibility, Homophily, and Attraction." In: *Journal of Computer-Mediated Communication* 32 (2005), pp. 178–197. ISSN: 0360-3989. DOI: `10.1111/j.1468-2958.2006.00008.x` (cit. on pp. 117, 120, 139).

[222] B. O. Olatunji, J. Haidt, D. McKay, B. David. "Core, animal reminder, and contamination disgust: Three kinds of disgust with distinct personality, behavioral, physiological, and clinical correlates." In: *Journal of Research in Personality* 42.5 (2008), pp. 1243–1259. ISSN: 0092-6566. DOI: `10.1016/j.jrp.2008.03.009` (cit. on p. 155).

[223] G. H. Orians. "An evolutionary perspective on aesthetics." In: *Bulletin of Psychology and the Arts* 2.1 (2001), pp. 25–29 (cit. on p. 87).

[224] L. A. Parr. "The evolution of face processing in primates." In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 366.1571 (2011), pp. 1764–1777. ISSN: 0962-8436. DOI: `10.1098/rstb.2010.0358` (cit. on p. 113).

[225] O. Pascalis, M. de Haan, C. A. Nelson. "Is Face Processing Species-Specific During the First Year of Life?" In: *Science* 296.5571 (2002), pp. 1321–1323. ISSN: 0036-8075. DOI: `10.1126/science.1070223` (cit. on p. 78).

[226] T. C. Peck, S. Seinfeld, S. M. Aglioti, M. Slater. "Putting yourself in the skin of a black avatar reduces implicit racial bias." In: *Consciousness and Cognition* 22.3 (2013), pp. 779–787. ISSN: 1053-8100. DOI: `10.1016/j.concog.2013.04.016` (cit. on pp. 187, 243).

[227] D. I. Perrett, D. Burt, I. S. Penton-Voak, K. J. Lee, D. a. Rowland, R. Edwards. "Symmetry and Human Facial Attractiveness." In: *Evolution and Human Behavior* 20 (1999), pp. 295–307. ISSN: 1090-5138. DOI: `10.1016/S1090-5138(99)00014-8` (cit. on pp. 86, 93).

[228] L. Pessoa. "On the relationship between emotion and cognition." In: *Nature Reviews Neuroscience* 9.2 (2008), pp. 148–158. ISSN: 1471-003X. DOI: `10.1038/nrn2317` (cit. on p. 86).

[229] I. V. Piryankova, H. Y. Wong, S. A. Linkenauger, C. Stinson, M. R. Longo, H. H. Bülthoff, B. J. Mohler. "Owning an Overweight or Underweight Body: Distinguishing the Physical, Experienced and Virtual Body." In: *PLOS ONE* 9.8 (2014), pp. 1–13. DOI: 10.1371/journal.pone.0103428 (cit. on p. 187).

[230] L. Piwek, L. S. McKay, F. E. Pollick. "Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion." In: *Cognition* 130.3 (2014), pp. 271–277. ISSN: 0010-0277. DOI: 10.1016/j.cognition.2013.11.001 (cit. on pp. 74, 89, 156).

[231] E. A. Poe. *Maelzel's Chess Player*. Southern Literary Messenger, 1836 (cit. on p. 59).

[232] E. Poliakoff, N. Beach, R. Best, T. Howard, E. Gowen. "Can looking at a hand make your skin crawl? Peering into the uncanny valley for hands." In: *Perception* 42.9 (2013), pp. 998–1000. ISSN: 0301-0066. DOI: 10.1068/p7569 (cit. on pp. 74, 91).

[233] F. E. Pollick. "In search of the uncanny valley." In: *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*. Vol. 40 LNICST. 2010, pp. 69–78. ISBN: 3-642-12629-4. DOI: 10.1007/978-3-642-12630-7_8 (cit. on pp. 33, 80).

[234] M. Pouliquen, A. Bernard, J. Marsot, L. Chodorge. "Virtual hands and virtual reality multimodal platform to design safer industrial systems." In: *Computers in Industry* 58.1 (2007), pp. 46–56. ISSN: 0166-3615. DOI: 10.1016/j.compind.2006.04.001 (cit. on p. 189).

[235] D. Powell. "Chimera contemporary: The enduring art of the composite beast." In: *Leonardo* 37.4 (2004), pp. 332–340. DOI: 10.1162/0024094041724472 (cit. on p. 151).

[236] S. D. Preston, F. B. M. de Waal. "Empathy: Its ultimate and proximate bases." In: *Behavioral and Brain Sciences* 25.1 (2002), 1–20. DOI: 10.1017/S0140525X02000018 (cit. on p. 85).

[237] E. P. Prokopakis, I. M. Vlastos, V. Picavet, G. Nolst, R. Thomas, C. Cingi, P. W. Hellings. "The golden ratio in facial symmetry." In: *Rhinology* 51.4 (2013), pp. 18–21. DOI: 10.4193/Rhino12.111 (cit. on p. 124).

[238] C. H. Ramey. “The uncanny valley of similarities concerning abortion, baldness, heaps of sand, and humanlike robots.” In: *Proceedings of Views of the Uncanny Valley Workshop: IEEE-RAS International Conference on Humanoid Robots*. 2005, pp. 8–13 (cit. on pp. 85, 86, 92, 154, 156).

[239] J. E. Raymond, M. J. Fenske, N. Westoby. “Emotional Devaluation of Distracting Patterns and Faces: A Consequence of Attentional Inhibition During Visual Search?” In: *Journal of Experimental Psychology: Human Perception and Performance* 31.6 (2005), pp. 1404–1415. ISSN: 1939-1277. DOI: 10.1037/0096-1523.31.6.1404 (cit. on pp. 81, 82).

[240] J. Reichardt. *Robots: Fact, Fiction, and Prediction*. London: Penguin Books, 1978, p. 168. ISBN: 978-0-140-04938-1 (cit. on pp. 20, 29, 33).

[241] T. Reid, G. Yaffe, R. Nichols. “Thomas Reid.” In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Fall 2014. 2014 (cit. on p. 68).

[242] J. Reilly. “ESRB Made Valve Change Left 4 Dead 2 Cover.” In: *IGN* (2009) (cit. on p. 187).

[243] P. Renaud, J. L. Rouleau, L. Granger, I. Barsetti, S. Bouchard. “Measuring sexual preferences in virtual reality: A pilot study.” In: *CyberPsychology & Behavior* 5.1 (2002), pp. 1–9. DOI: 10.1089/109493102753685836 (cit. on p. 195).

[244] B. H. Repp. “Categorical Perception: Issues, Methods, Findings.” In: *Speech and Language* 10 (1984), pp. 243–335. ISSN: 0193-3434. DOI: 10.1016/B978-0-12-608610-2.50012-1 (cit. on p. 76).

[245] G. Rhodes. “The Evolutionary Psychology of Facial Beauty.” In: *Annual Review of Psychology* 57.1 (2006), pp. 199–226. ISSN: 0066-4308. DOI: 10.1146/annurev.psych.57.102904.190208 (cit. on pp. 86, 93).

[246] M. Rice, R. Koh, Q. Lui, Q. He, M. Wan, V. Yeo, J. Ng, W. P. Tan. “Comparing Avatar Game Representation Preferences Across Three Age Groups.” In: *Proceedings of the 2013 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’13. Paris, France, 2013, pp. 1161–1166. ISBN: 978-1-4503-1952-2. DOI: 10.1145/2468356.2468564 (cit. on p. 118).

[247] S. Rieckhoff-Pauli, J. Biel, B. Abels. *Die Kelten in Deutschland*. Theiss, 2001. ISBN: 978-3-806-21367-6 (cit. on p. 19).

[248] G. Rizzolatti, L. Fogassi, V. Gallese. “Neurophysiological mechanisms underlying the understanding and imitation of action.” In: *Nature Reviews Neuroscience* 2.9 (2001), pp. 661–670. ISSN: 1471-0048. DOI: 10.1038/35090060 (cit. on p. 85).

[249] T. Ro, A. Friggel, N. Lavie. “Attentional biases for faces and body parts.” In: *Visual Cognition* 15.3 (2007), pp. 322–348. ISSN: 1350-6285. DOI: 10.1080/13506280600590434 (cit. on p. 97).

[250] D. Roberson, J. Davidoff. “The categorical perception of colors and facial expressions: The effect of verbal interference.” In: *Memory & Cognition* 28.6 (2000), 977—986. ISSN: 0090-502X. DOI: 10.3758/bf03209345 (cit. on p. 76).

[251] E. Rolls. *Emotion and Decision-Making Explained*. OUP Oxford, 2013. ISBN: 978-0-191-63514-4 (cit. on p. 86).

[252] S. Rose. *Tintin and the Uncanny Valley: when CGI gets too real*. The Guardian. 2011 (cit. on p. 66).

[253] R. Rosenthal, L. Jacobson. “Pygmalion in the classroom.” In: *The Urban Review* 3.1 (1968), pp. 16–20. ISSN: 0042-0972. DOI: 10.1007/BF02322211 (cit. on p. 50).

[254] M. E. Rosheim. “In the footsteps of Leonardo (articulated anthropomorphic robot).” In: *IEEE Robotics & Automation Magazine* 4.2 (1997), pp. 12–14. ISSN: 1070-9932 (cit. on p. 56).

[255] J.-J. Rousseau, H. Coignet, J. Waeber. *Pygmalion: scène lyrique*. Editions Université-Conservatoire de musique, 1997. ISBN: 978-2-884-33010-7 (cit. on p. 50).

[256] P. Rozin, A. E. Fallon. “A perspective on disgust.” In: *Psychological Review* 94.1 (1987), pp. 23–41. ISSN: 1939-1471. DOI: 10.1037/0033-295X.94.1.23 (cit. on p. 83).

[257] M. Ruckenstein. “Toying with the world: Children, virtual pets and the value of mobility.” In: *Childhood* 17.4 (2010), pp. 500–513. ISSN: 0907-5682. DOI: 10.1177/0907568209352812 (cit. on p. 153).

[258] K. Rujirakul, C. So-In, B. Arnonkijpanich. “PEM-PCA: A parallel expectation-maximization PCA face recognition architecture.” In: *The Scientific World Journal* 2014 (2014). ISSN: 1537-744X. DOI: 10.1155/2014/468176 (cit. on p. 140).

[259] J. A. Russell. “Core Affect and the Psychological Construction of Emotion.” In: *Psychological Review* 110.1 (2003), pp. 145–172. DOI: 10.1037/0033-295X.110.1.145 (cit. on p. 73).

[260] S. Ryan, M. Oaten, R. J. Stevenson, T. I. Case. “Facial disfigurement is treated like an infectious disease.” In: *Evolution and Human Behavior* 33.6 (2012), pp. 639–646. ISSN: 1090-5138. DOI: 10.1016/j.evolhumbehav.2012.04.001 (cit. on p. 83).

[261] H. Sakaguchi, M. Sakakibara, A. Reinert, A. Baldwin, S. Buscemi, Ming-Na, V. Rhames. *Final Fantasy: The Spirits Within*. Colombia Pictures. 2001 (cit. on pp. 29, 67).

[262] A. P. Saygin, T. Chaminade, H. Ishiguro. "The Perception of Humans and Robots: Uncanny Hills in Parietal Cortex." In: *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* May (2009), pp. 2004–2008 (cit. on p. 85).

[263] A. P. Saygin, T. Chaminade, H. Ishiguro, J. Driver, C. Frith. "The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions." In: *Social Cognitive and Affective Neuroscience* 7.4 (2012), pp. 413–422. ISSN: 1749-5016. DOI: 10.1093/scan/nsr025 (cit. on p. 89).

[264] M. Schaller, J. H. Park. "The Behavioral Immune System (and Why It Matters)." In: *Current Directions in Psychological Science* 20.2 (2011), pp. 99–103. ISSN: 0963-7214. DOI: 10.1177/0963721411402596 (cit. on p. 83).

[265] M. Schaller, G. E. Miller, W. M. Gervais, S. Yager, E. Chen. "Mere Visual Perception of Other People's Disease Symptoms Facilitates a More Aggressive Immune Response." In: *Psychological Science* 21.5 (2010), pp. 649–652. ISSN: 0956-7976. DOI: 10.1177/0956797610368064 (cit. on p. 83).

[266] M. Schmidt, J. X. Kafka, O. D. Kothgassner, H. Hlavacs, L. Beutl, A. Felnhofer. "Why Does It Always Rain on Me? Influence of Gender and Environmental Factors on Usability, Technology Related Anxiety and Immersion in Virtual Environments." In: *Advances in Computer Entertainment: 10th International Conference, ACE 2013, Boekelo, The Netherlands, November 12-15, 2013. Proceedings*. Ed. by D. Reidsma, H. Katayose, A. Nijholt. Cham: Springer International Publishing, 2013, pp. 392–402. ISBN: 978-3-319-03161-3. DOI: 10.1007/978-3-319-03161-3_29 (cit. on pp. 187, 238).

[267] E. Schneider, Y. Wang, S. Yang. "Exploring the Uncanny Valley with Japanese Video Game Characters." In: *Proceedings of the DiGRA International Conference: Situated Play*. DIGRA '07. 2007, pp. 546–549 (cit. on p. 90).

[268] J. R. Schoenherr, T. J. Burleigh. "Uncanny Sociocultural Categories." In: *Frontiers in Psychology* 5.January (2015), pp. 1–4. DOI: 10.3389/fpsyg.2014.01456 (cit. on p. 30).

[269] R. Schroeder. *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments*. Computer Supported Cooperative Work. Springer London, 2012. ISBN: 978-1-447-10277-9 (cit. on p. 21).

[270] T. Schubert, F. Friedmann, H. Regenbrecht. "The Experience of Presence: Factor Analytic Insights." In: *Presence: Teleoperators and Virtual Environments* 10.3 (2001), pp. 266–281. ISSN: 1054-7460. DOI: 10.1162/105474601300343603 (cit. on p. 186).

[271] G. Schwab. *Sagen des klassischen Altertums*. Anaconda Verlag, 1975, 96 figs. ISBN: 978-3-866-47687-5 (cit. on p. 48).

[272] V. Schwind. "Historical, Cultural, and Aesthetic Aspects of the Uncanny Valley." In: *Collective Agency and Cooperation in Natural and Artificial Systems*. Cham: Springer International Publishing, 2015. Chap. 5, pp. 81–107. DOI: 10.1007/978-3-319-15515-9_5 (cit. on pp. 39, 45, 96).

[273] V. Schwind, S. Jäger. "The Uncanny Valley and the Importance of Eye Contact." In: *Mensch und Computer 2015 - Tagungsband*. Vol. 15. Berlin, München, Boston: DeGruyter, 2015, pp. 153–162. DOI: 10.1515/9783110443929-017 (cit. on pp. 39, 95).

[274] V. Schwind, S. Jäger. "The Uncanny Valley and the Importance of Eye Contact." In: *i-com* 15.1 (2016), pp. 93–104. ISSN: 2196-6826. DOI: 10.1515/icom-2016-0001 (cit. on pp. 39, 95, 181).

[275] V. Schwind, N. Pohl, P. Bader. "Accuracy of a Low-Cost 3D-printed Head-Mounted Eye Tracker." In: *Mensch und Computer 2015 - Tagungsband*. Berlin, München, Boston: DeGruyter, 2015, pp. 259–262. DOI: 10.1515/9783110443929-028 (cit. on pp. 80, 102).

[276] V. Schwind, K. Wolf, N. Henze. "Design Guidelines for Virtual Characters to avoid the Uncanny Valley." In: *interactions* (2018). DOI: 10.1145/3236673 (cit. on p. 245).

[277] V. Schwind, K. Wolf, N. Henze. "FaceMaker—A Procedural Face Generator to Foster Character Design Research." In: *Game Dynamics*. Cham: Springer International Publishing, 2017, pp. 95–113. ISBN: 978-3-319-53087-1. DOI: 10.1007/978-3-319-53088-8_6 (cit. on pp. 40, 42, 115).

[278] V. Schwind, K. Wolf, N. Henze, O. Korn. “Determining the Characteristics of Preferred Virtual Faces Using an Avatar Generator.” In: *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY ’15. London, United Kingdom: ACM, 2015, pp. 221–230. ISBN: 978-1-4503-3466-2. DOI: `10.1145/2793107.2793116` (cit. on pp. 40, 87, 115).

[279] V. Schwind, K. Leicht, S. Jäger, K. Wolf, N. Henze. “Is there an Uncanny Valley of Virtual Animals? A Quantitative and Qualitative Investigation.” In: *International Journal of Human-Computer Studies* 111 (2018), pp. 49 –61. ISSN: 1071-5819. DOI: `https://doi.org/10.1016/j.ijhcs.2017.11.003` (cit. on pp. 40, 149, 254).

[280] V. Schwind, J. Reinhardt, R. Rzayev, N. Henze, K. Wolf. “On the Need for Standardized Methods to Study the Social Acceptability of Emerging Technologies.” 2018 (cit. on p. 42).

[281] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, N. Henze. “"These Are Not My Hands!": Effect of Gender on the Perception of Avatar Hands in Virtual Reality.” In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17. Denver, CO, USA: ACM, 2017, pp. 1577–1582. ISBN: 978-1-4503-4655-9. DOI: `10.1145/3025453.3025602` (cit. on pp. 40, 185, 210, 254).

[282] V. Schwind, L. Lin, M. Di Luca, S. Jörg, J. Hillis. “Touch with Foreign Hands: The Effect of Virtual Hand Appearance on Visual-Haptic Integration.” In: *Symposium of Applied Perception*. ACM, 2017. DOI: `10.1145/3225153.3225158` (cit. on p. 42).

[283] V. Schwind, P. Bader, S. Schneegass, K. Wolf, N. Henze. “Towards Transparent Handheld See-Through Devices.” In: *NordiCHI 2014 workshop on Interactions and Applications on See-through Technologies*. 2014 (cit. on p. 42).

[284] V. Schwind, J. Reinhardt, R. Rzayev, N. Henze, K. Wolf. “Virtual Reality on the Go ? A Study on Social Acceptance of VR Glasses.” In: *Adjunct proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM Press, Jan. 1, 2018. ISBN: 9781450359412. DOI: `10.1145/3236112.3236127`. published (cit. on p. 42).

[285] V. Schwind, P. Knierim, L. Chuang, N. Henze. “"Where’s Pinky?": The Effects of a Reduced Number of Fingers in Virtual Reality.” In: *Proceedings of the 2017*

CHIPLAY Conference on Computer-Human Interaction in Play. CHI PLAY '17. Amsterdam, The Netherlands: ACM, 2017, pp. 507–515. ISBN: 978-1-4503-4898-0. DOI: 10.1145/3116595.3116596 (cit. on pp. 40, 185, 254).

[286] R. Scott. *Blade Runner - The Director's Cut*. Warner Bros. 1992 (cit. on pp. 65, 66).

[287] J. Seyama, R. S. Nagayama. "The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces." In: *Presence: Teleoperators and Virtual Environments* 16.4 (2007), pp. 337–351. ISSN: 1054-7460. DOI: 10.1162/pres.16.4.337 (cit. on pp. 74, 80, 81, 93, 144, 150, 157, 172, 177, 180, 181).

[288] W. Shebilske, D. Fisher. *Understanding extended discourse through the eyes: How and why*. Lawrence Erlbaum Associates: Hillsdale, NJ. 1983 (cit. on p. 96).

[289] M. W. Shelley. *Frankenstein or The Modern Prometheus*. Lackington Hughes Harding Mavor & Jones, 1818, p. 280 (cit. on p. 61).

[290] B. Shivali. *Meet China's freakishly realistic droids: Creepy machines go on display at the World Robot conference*. MailOnline. 2016 (cit. on p. 32).

[291] E. A. Simpson, K. Varga, J. E. Frick, D. Fragaszy. "Infants Experience Perceptual Narrowing for Nonprimate Faces." In: *Infancy* 16.3 (2011), pp. 318–328. ISSN: 1525-0008. DOI: 10.1111/j.1532-7078.2010.00052.x (cit. on p. 78).

[292] P. W. Singer. *Wired for War: The Robotics Revolution and Conflict in the 21st Century*. Vol. 23. A Penguin Book. Technology/Military Science. Penguin Press, 2009, pp. 312–313. ISBN: 1-594-20198-6. DOI: 10.1111/j.1747-7093.2009.00222_4.x (cit. on pp. 70, 85).

[293] M. Slater, B. Spanlang, M. V. Sanchez-Vives, O. Blanke. "First Person Experience of Body Transfer in Virtual Reality." In: *PLOS ONE* 5.5 (2010), pp. 1–9. DOI: 10.1371/journal.pone.0010564 (cit. on pp. 188, 190).

[294] R. W. Soukoreff, I. S. MacKenzie. "Metrics for Text Entry Research: An Evaluation of MSD and KSPC, and a New Unified Error Metric." In: *Proceedings of the 2003 CHI Conference on Human Factors in Computing Systems*. CHI '03. Ft. Lauderdale, Florida, USA: ACM, 2003, pp. 113–120. ISBN: 1-58113-630-7. DOI: 10.1145/642611.642632 (cit. on pp. 228, 231).

[295] S. Spielberg. *The Adventures of Tintin*. Paramount Pictures (North America), Columbia Pictures (International). 2011 (cit. on p. 67).

[296] K. M. Stanney, K. S. Kingdon, D. Graeber, R. S. Kennedy. “Human Performance in Immersive Virtual Environments: Effects of Exposure Duration, User Control, and Scene Complexity.” In: *Human Performance* 15.4 (2002), pp. 339–366. DOI: 10.1207/S15327043HUP1504_03 (cit. on p. 195).

[297] S. a. Steckenfinger, A. a. Ghazanfar. “Monkey visual behavior falls into the uncanny valley.” In: *Proceedings of the National Academy of Sciences of the United States of America* 106.43 (2009), pp. 18362–6. ISSN: 1091-6490. DOI: 10.1073/pnas.0910063106 (cit. on pp. 67, 83, 92, 96, 97, 112, 155, 156, 181).

[298] J.-P. Stein, P. Ohler. “Venturing into the uncanny valley of mind—The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting.” In: *Cognition* 160.Supplement C (2017), pp. 43 –50. ISSN: 0010-0277. DOI: https://doi.org/10.1016/j.cognition.2016.12.010 (cit. on pp. 91, 186).

[299] W. Steptoe, A. Steed, M. Slater. “Human Tails: Ownership and Control of Extended Humanoid Avatars.” In: *IEEE Transactions on Visualization and Computer Graphics* 19.4 (2013), pp. 583–590. ISSN: 1077-2626. DOI: 10.1109/TVCG.2013.32 (cit. on p. 190).

[300] M. K. Strait, V. A. Floerke, W. Ju, K. Maddox, J. D. Remedios, M. F. Jung, H. L. Urry. “Understanding the Uncanny: Both Atypical Features and Category Ambiguity Provoke Aversion toward Humanlike Robots.” In: *Frontiers in Psychology* 8 (2017). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2017.01366 (cit. on pp. 81, 87, 93).

[301] M. Strait, L. Vujovic, V. Floerke, M. Scheutz, H. Urry. “Too Much Humanness for Human-Robot Interaction.” In: *Proceedings of the 2015 CHI Conference on Human Factors in Computing Systems*. CHI ’15. New York, NY, USA: ACM Press, 2015, pp. 3593–3602. ISBN: 978-1-450-33145-6. DOI: 10.1145/2702123.2702415 (cit. on pp. 87, 90).

[302] S. Strandh. *Die Maschine: Geschichte, Elemente, Funktion Ein enzyklopädisches Sachbuch*. Weltbild-Verlag, 1992. ISBN: 978-3-893-50052-9 (cit. on pp. 19, 55–58).

[303] R. Styles. *Secrets of men who dress up as rubber dolls revealed in new documentary*. dailymail.co.uk. 2014 (cit. on p. 50).

[304] J. P. Swaddle, I. G. Guthill. “Asymmetry and human facial attractiveness: symmetry may not always be beautiful.” In: 261.1360 (1995), pp. 111–116. ISSN: 0962-8452. DOI: 10.1098/rspb.1995.0124 (cit. on p. 139).

[305] N. Sweeney. *The Secret Of The Living Dolls*. 4OD - Channel 4. 2014 (cit. on p. 50).

[306] E. Thiele. *Karel Čapek*. Vol. 1257. P. Reclam, 1988 (cit. on p. 63).

[307] F. Thomas, O. Johnston, F. Thomas. *The illusion of life: Disney animation*. Hyperion New York, 1995 (cit. on p. 20).

[308] J. Thompson, G. Trafton, M. McCurry, E. Francis. "Perceptions of an animated figure as a function of movement naturalness: No sign of the uncanny valley." In: *Journal of Vision* 8.6 (2010), pp. 906–906. ISSN: 1534-7362. DOI: 10.1167/8.6.906 (cit. on pp. 30, 89, 156).

[309] R. Thornhill, S. W. Gangestad. "Human facial beauty - Averageness, symmetry, and parasite resistance." In: *Human Nature* 4.3 (1993), pp. 237–269. ISSN: 1045-6767. DOI: 10.1007/BF02692201 (cit. on pp. 86, 87, 93).

[310] W. H. Thorpe. "Some Problems of Animal Learning." In: *Proceedings of the Linnean Society of London* 156.2 (1944), pp. 70–83. ISSN: 0370-0461. DOI: 10.1111/j.1095-8312.1944.tb00374.x (cit. on p. 50).

[311] P. Throop. *Hildegard von Bingen's Physica: The Complete English Translation of Her Classic Work on Health and Healing*. Inner Traditions/Bear, 1998. ISBN: 978-1-594-77775-2 (cit. on p. 55).

[312] B. Tiddeman, M. Burt, D. Perrett. "Prototyping and transforming facial textures for perception research." In: *IEEE Computer Graphics and Applications* 21.October (5 2001), pp. 42–50. ISSN: 0272-1716. DOI: 10.1109/38.946630 (cit. on p. 123).

[313] A. Tinwell, M. Grimshaw, D Abdel-Nabi. "The uncanny valley and nonverbal communication in virtual characters." In: *Nonverbal Communication in Virtual Worlds* (2014), pp. 325–342 (cit. on p. 90).

[314] A. Tinwell, M. Grimshaw, D. Abdel-Nabi. "Effect of emotion and articulation of speech on the uncanny valley in virtual characters." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 6975 LNCS. Springer, 2011, pp. 557–566. ISBN: 978-3-642-24570-1. DOI: 10.1007/978-3-642-24571-8_69 (cit. on p. 67).

[315] A. Tinwell, M. Grimshaw, D. A. Nabi. "The Effect of Onset Asynchrony in Audio Visual Speech and the Uncanny Valley in Virtual Characters." In: *International Journal of the Digital Human* (2014). DOI: 10.1504/IJMRS.2015.068991 (cit. on p. 90).

[316] A. Tinwell, M. Grimshaw, A. Williams. “Uncanny behaviour in survival horror games.” In: *Journal of Gaming & Virtual Worlds* 2.1 (2010), pp. 3–25. ISSN: 1757-191X. DOI: 10.1386/jgvw.2.1.3_1 (cit. on pp. 97, 122).

[317] A. Tinwell, D. A. Nabi, J. P. Charlton. “Perception of psychopathy and the Uncanny Valley in virtual characters.” In: *Computers in Human Behavior* 29 (2013), pp. 1617–1625. ISSN: 0747-5632. DOI: 10.1016/j.chb.2013.01.008 (cit. on pp. 90, 96).

[318] A. Tinwell, M. Grimshaw, D. A. Nabi, A. Williams. “Facial expression of emotion and perception of the Uncanny Valley in virtual characters.” In: *Computers in Human Behavior* 27 (2011), pp. 741–749. ISSN: 0747-5632. DOI: 10.1016/j.chb.2010.10.018 (cit. on pp. 90, 92, 156, 157, 169, 172, 178).

[319] B. A. Urgen, A. X. Li, C. Berka, M. Kutas, H. Ishiguro, A. P. Saygin. “Predictive coding and the Uncanny Valley hypothesis: Evidence from electrical brain activity.” In: *Proceedings of the 10th ACM/IEEE Internation Conference on Human-Robot Interaction* (2015) (cit. on p. 89).

[320] V. C. Valdes, J. L. Bischoff. “Accelerator 14C dates for early upper paleolithic (basal Aurignacian) at El Castillo Cave (Spain).” In: *Journal of Archaeological Science* 16.6 (1989), pp. 577–584. ISSN: 0305-4403. DOI: 10.1016/0305-4403(89)90023-X (cit. on p. 18).

[321] S. Valverde. “The Modern Sex Doll-Owner: A Descriptive Analysis.” MS in Psychology. Cal Poly State University, 2012 (cit. on p. 50).

[322] G. van Belle, M. Ramon, P. Lefèvre, B. Rossion. “Fixation patterns during recognition of personally familiar and unfamiliar faces.” In: *Frontiers in Psychology* 1.June (2010), p. 20. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2010.00020 (cit. on p. 96).

[323] C. Velocci. *’Alita: Battle Angel’ Doesn’t Understand Why Anime Characters Have Big Eyes*. waypoint.vice.com. 2017 (cit. on p. 80).

[324] V. Vinayagamoorthy, A. Brogni, M. Gillies, M. Slater, A. Steed. “An investigation of presence response across variations in visual realism.” In: *The 7th Annual International Presence Workshop*. 2004, pp. 148–155 (cit. on pp. 91, 186, 204).

[325] C. Vinette, F. Gosselin, P. G. Schyns. “Spatio-temporal dynamics of face recognition in a flash: It’s in the eyes.” In: *Cognitive Science* 28.2 (2004), pp. 289–301. ISSN: 0364-0213. DOI: 10.1016/j.cogsci.2004.01.002 (cit. on p. 97).

[326] A. Voit, T. Machulla, D. Weber, V. Schwind, S. Schneegass, N. Henze. “Exploring notifications in smart home environments.” In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI ’16. New York, NY, USA: ACM Press, 2016, pp. 942–947. ISBN: 978-1-450-34413-5. DOI: 10.1145/2957265.2962661 (cit. on p. 42).

[327] E. Voland, K. Grammer. *Evolutionary Aesthetics*. Springer Berlin Heidelberg, 2013. ISBN: 978-3-662-07142-7 (cit. on p. 88).

[328] K. Völker. *Künstliche Menschen: Dichtungen und Dokumente über Golems, Homunculi, lebende Statuen und Androiden*. Phantastische Bibliothek. Suhrkamp Verlag KG, 1994, p. 512. ISBN: 978-3-518-38793-1 (cit. on pp. 53, 55, 58).

[329] A. Vosskühler, V. Nordmeier, L. Kuchinke, A. M. Jacobs. “OGAMA (Open Gaze and Mouse Analyzer): open-source software designed to analyze eye and mouse movements in slideshow study designs.” In: *Behavior research methods* 40.4 (2008), pp. 1150–1162. ISSN: 1554-351X. DOI: 10.3758/BRM.40.4.1150 (cit. on p. 100).

[330] A. Wachowski. *The Matrix*. Warner Bros. Pictures, Roadshow Entertainment. 1999 (cit. on p. 67).

[331] J. Walker, B. Li, K. Vertanen, S. Kuhl. “Efficient Typing on a Visually Occluded Physical Keyboard.” In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17. Denver, CO, USA: ACM, 2017, pp. 5457–5461. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025783 (cit. on p. 235).

[332] H. G. Wells. *Mr. Wells reviews a current Film: He Takes Issue With This German Conception of What The City of One Hundred Years Hence Will Be Like*. The New York Times Company. New York, NY, USA, 1927 (cit. on p. 65).

[333] S. Wells. *Mars Needs Moms!* Walt Disney Pictures. 2011 (cit. on pp. 31, 67).

[334] L. Weschler. “Wired 10.06:Why Is This Man Smiling?” In: *Wired* (2006), pp. 1–6 (cit. on p. 31).

[335] C. Westbrook. *Secrets Of The Living Dolls pulls in and freaks out viewers in equal measure*. metro.co.uk. 2014 (cit. on p. 50).

[336] B. G. Witmer, M. J. Singer. “Measuring Presence in Virtual Environments: A Presence Questionnaire.” In: *Presence: Teleoperators and Virtual Environments* 7.3 (1998), pp. 225–240. ISSN: 1054-7460. DOI: 10.1162/105474698565686 (cit. on pp. 194, 204, 207, 227, 235).

[337] J. O. Wobbrock, L. Findlater, D. Gergle, J. J. Higgins. “The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures.” In: *Proceedings of the 2011 CHI Conference on Human Factors in Computing Systems*. CHI ’11. Vancouver, BC, Canada, 2011, pp. 143–146. ISBN: 978-1-450-30267-8. DOI: 10.1145/1978942.1978963 (cit. on pp. 197, 209).

[338] M. Woerner. *"New Tron Legacy trailer shows off Jeff Bridges’ CG evil twin"*. IO9. 2010 (cit. on p. 31).

[339] K. Wolf, S. Schneegass, N. Henze, D. Weber, V. Schwind, P. Knierim, S. Mayer, T. Dingler, Y. Abdelrahman, T. Kubitza, M. Funk, A. Mebus, A. Schmidt. “TUIs in the Large: Using Paper Tangibles with Mobile Devices.” In: *Proceedings of the 2015 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’15. New York, NY: ACM, 2015, pp. 1579–1584. DOI: 10.1145/2702613.2732863 (cit. on p. 42).

[340] G. Wood. *Living Dolls: A Magical History of the Quest for Mechanical Life*. Faber & Faber, 2002. ISBN: 978-0-571-21466-2 (cit. on p. 19).

[341] M. Wrzesien, J.-M. Burkhardt, C. Botella, M. Alcañiz. “Towards a Virtual Reality- and Augmented Reality-Mediated Therapeutic Process model: a theoretical revision of clinical issues and HCI issues.” In: *Theoretical Issues in Ergonomics Science* 16.2 (2015), pp. 124–153. ISSN: 1463-922X. DOI: 10.1080/1463922X.2014.903307 (cit. on p. 153).

[342] Y. Yamada, T. Kawabe, K. Ihaya. “Categorization difficulty is associated with negative evaluation in the "uncanny valley" phenomenon.” In: *Japanese Psychological Research* 55 (2013), pp. 20–32. ISSN: 0021-5368. DOI: 10.1111/j.1468-5884.2012.00538.x (cit. on pp. 74, 76, 77, 82, 84, 86, 154, 171).

[343] C. Youngblut, O. Huie. “The relationship between presence and performance in virtual environments: results of a VERTS study.” In: *IEEE Virtual Reality*. Vol. 2003-Janua. IEEE VR ’03. IEEE Computational Society, 2003, pp. 277–278. ISBN: 0-7695-1882-6. DOI: 10.1109/VR.2003.1191158 (cit. on p. 195).

[344] C. Zanbaka, P. Goolkasian, L. Hodges. “Can a Virtual Cat Persuade You?: The Role of Gender and Realism in Speaker Persuasiveness.” In: *Proceedings of the 2006 CHI Conference on Human Factors in Computing Systems*. CHI ’06. Montréal, Québec, Canada: ACM, 2006, pp. 1153–1162. ISBN: 1-59593-372-7. DOI: 10.1145/1124772.1124945 (cit. on p. 157).

[345] E. Zell, C. Aliaga, A. Jarabo, K. Zibrek, D. Gutierrez, R. McDonnell, M. Botsch. "To Stylize or not to Stylize? The Effect of Shape and Material Stylization on the Perception of Computer-Generated Faces." In: *ACM Transactions on Graphics* 34.6 (2015), p. 12. ISSN: 0730-0301. DOI: 10.1145/2816795.2818126 (cit. on pp. 87, 182).

[346] R. Zemeckis. *The Polar Express*. Warner Bros. Pictures/Castle Rock Entertainment/Shangri-La Entertainment, ImageMovers/Playtone/Golden Mean. 2004 (cit. on pp. 29, 67).

[347] J. Zlatev. "The epigenesis of meaning in human beings, and possibly in robots." In: *Minds and Machines* 11 (2001), pp. 155–195. ISSN: 0924-6495. DOI: 10.1023/A:1011218919464 (cit. on p. 68).

[348] J. A. Złotowski, H. Sumioka, S. Nishio, D. F. Glas, C. Bartneck, H. Ishiguro. "Persistence of the uncanny valley: the influence of repeated interactions and a robot's attitude on its perception." In: *Frontiers in Psychology* 6 (2015), pp. 1–13. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2015.00883 (cit. on p. 89).

All URLs cited were checked in 2018-01-12.

List of Figures

List of Tables

List of Acronyms

AI artificial intelligence

AOI area of interest

AR augmented reality

CG computer-generated

CGI computer-generated imagery

CV 1 Consumer Version 1

EEG electroencephalography

ERP event-related potentials

fMRI functional magnetic resonance imaging

FoV field of view

FPS frames per second

HCI human-computer interaction

HdM Hochschule der Medien

HQ high-quality

HRI human-robot interaction

MuC Mensch und Computer

NPC non-player character

PC personal computer

PQ presence questionnaire

RM repeated measure

RPG role-playing game

RQ research question

TCT task completion time

VFX visual effects

VR virtual reality

WPM words per minute