From Perception to Interaction with Virtual Characters

Eduard Zell[1], Katja Zibrek[1], Xueni Pan[2], Marco Gillies[2] and Rachel McDonnell[1]

[1]Trinity College Dublin
[2]Goldsmiths, University of London

Abstract

This course will introduce students, researchers and digital artists to the recent results in perceptual research on virtual characters. It covers how technical and artistic aspects that constitute the appearance of a virtual character influence human perception, and how to create a plausibility illusion in interactive scenarios with virtual characters. We will report results of studies that addressed the influence of low-level cues like facial proportions, shading or level of detail and higher-level cues such as behavior or artistic stylization. We will place emphasis on aspects that are encountered during character development, animation, interaction design and achieving consistency between the visuals and storytelling. We will close with the relationship between verbal and non-verbal interaction and introduce some concepts which are important for creating convincing character behavior in virtual reality. The insights that we present in this course will serve as an additional toolset to anticipate the effect of certain design decisions and to create more convincing characters, especially in the case where budgets or time are limited.

1. Course Description

Virtual humans are finding a growing number of applications, such as in social media apps, Spaces by Facebook, Bitmoji and Genies, as well as computer games and human-computer interfaces. Their use today has also extended from the typical on-screen display applications to immersive and collaborative environments (VR/AR/MR). At the same time, we are also witnessing significant improvements in real-time performance, increased visual fidelity of characters and novel devices. The question of how these developments will be received from the user's point of view, or which aspects of virtual characters influence the user more, has therefore never been so important. This course will provide an overview of existing perceptual studies related to the topic of virtual characters.

To make the course easier to follow, we start with a brief overview of human perception and how perceptual studies are conducted in terms of methods and experiment design. With knowledge of the methods, we continue with artistic and technical aspects which influence the design of character appearance (lighting and shading, facial feature placement, stylization, etc.). Important questions on character design will be addressed such as – if I want my character to be highly appealing, should I render with realistic or stylized shading? What facial features make my character appear more trustworthy? Do dark shadows enhance the emotion my character is portraying?

We then dive deeper into the movement of the characters, exploring which information is present in the motion cues and how motion can, in combination with character appearance, guide our perception and even be a foundation of biased perception (stereotypes). Some examples of questions that we will address are – if I want my character to appear extroverted, what movement or appearance is needed to achieve this? Can character appearance influence my moral decisions in a video game? We then start to make our way into the domain of virtual reality and how it can be used to study perception of virtual characters and explore how appearance of virtual characters could affect our empathy level towards them. We also discuss possible behavioral measures for studying perception in virtual reality (VR).

In the last Section, we focus on the question – How should we design interactions with virtual characters that improve task performance and are more immersive? Plausibility illusion is an important element in VR – it makes the VR experience more immersive, engaging, and ensures that skills learnt in VR can be directly applied to real life experiences. Starting with a brief review on publications evaluating plausibility illusion we will focus on the context of virtual characters, social presence or co-presence. The theory of plausibility illusion implies that the experience of interaction with a virtual character should be as close as possible to face-to-face interaction with a real person. Human face-to-face interaction is highly multimodal: the verbal content of conversation is enhanced by other, non-verbal signals that carry a lot of information, for example, tone of voice, facial expressions, gestures, gaze and spatial behaviour. Interaction with a character involves a tight loop of sensing a person and responses from the character. This course will cover sensing technologies, types of response and methods to map between the two. We will also address the relationship between verbal and non-verbal interaction including different roles that people adopt in conversation: speaking, listening, and other forms of non-

DOI: 10.2312/egt.20201001

verbal interaction. All of these issues will be informed by both the psychology of social interaction and current VR technology. We will use two examples to illustrate the design process of virtual character interaction in VR: one on training for doctor-patient communication, another one on our recent project with a game company on creating AI-characters for the '*Peaky Blinders*' VR game.

The course provides the overview of the relevant studies in a way that makes it easy to identify answers to practical questions in production and character development. At the same time, we avoid giving definite answers to questions of character and interaction design and encourage further investigation by listing questions left unanswered to allow for critical evaluation of the presented research.

Finally, participating in a perceptual experiment is a multi-modal experience, which cannot be reproduced only by descriptive reports of the experiment design. For this reason, we will select a few representative experiments and run a highly compact version of them during the course for illustration purposes. The stimuli will be shown on the projector wall and the participants will be able to rate the stimuli within a small time-frame using their smartphones. Experiments will primarily be selected to introduce a new topic. We are fully aware that the obtained results are not representative by any means, but we believe that such live surveys will improve the understanding of the study design, increase engagement of participants, and be a welcoming break during a 180 minute talk.

Previous Version of this Tutorial

At SIGGRAPH 2019, we gave a shorter version (90min) of this tutorial, which was well attended (approximately 100 participants). Participants, who were less familiar with character perception research, were especially positive about the applicability of the knowledge. Given the positive feedback, we extended the tutorial on the topic of interaction.

Other Related Tutorials and Courses

Courses of the last 10–15 years at SIGGRAPH, SIGGRAPH Asia or Eurographics covered topics such as experiment design [CW13], visual perception of simple 3D shapes [FS09], as well as perception in graphics with applications to display technologies and virtual environments [GCL*06a, TOY*07, MR08]. Other courses covered a mixture of low-level stimuli perception and application in graphics, where character perception was partly addressed as well [OHM*04, MMG11]. Finally, there are courses that focused on perception of specific aspects of virtual characters; these include: (i) the expressiveness of body motion [VGS*06, HOP09], (ii) crowds [BKA*14, HLLO10, DMTPT09, TOY*07] (iii) a multi-disciplinary study of emotions covering aspects of philosophy, psychology and physiology [Ges12] and (iv) the creation of believable characters for dialogues [JKF*11].

Our course is the first to cover perception of virtual humans in a single resource, and addresses much more recent work than previous courses. We feel that it will be accessible for non-experts and a starting point for further investigation on related topics.

Audience

This course is suited for students, who want to get an overview of recent developments of perceptual research on virtual characters and identify open topics. Furthermore, this course is particularly designed for researchers and artists who work on virtual characters but are less familiar with the perceptual research.

Prerequisites

Fundamentals about creating and animating virtual characters and knowledge about design and analysis of perception experiments is beneficial, but not required.

Difficulty

Beginner to Intermediate

Duration

2×90min

Tutorial Website

www.eduardzell.com/VirtualCharacters

2. Schedule

I **Visual Perception Basics** (20 minutes)

a. Perception
b. Experiment Design & Statistics
c. Stimuli Creation

II **Character Appearance** (40 minutes)

a. Character Stylization
b. Character Realism
c. Facial Proportions
d. Level of Detail
e. Skin Appearance
f. Lighting and Shading
g. Visual Attention of Facial and Body Parts

III **Character Motion and Behaviour** (40 minutes)

a. Emotion
b. Gender
c. Gender and Emotion Bias
d. Personality
e. From perception to interaction

IV **Character Interaction** (80 minutes)

a. Plausibility Illusion
b. User input in VR
c. Social interaction and non-verbal behaviour
d. Case Studies

3. Lecturer Biographies

Eduard Zell

Eduard Zell is a postdoctoral Research Fellow at Trinity College Dublin. In 2018, he received his PhD from Bielefeld University, Germany on the topic of creation, animation and perception of virtual faces. His work was published at SIGGRAPH and Scientific Reports and for his thesis, he received the best thesis of the faculty award as well as the Eurographics PhD Award. Prior to his PhD, he completed a highly practical degree in Computer Animation and Visual Effects (M.Sc.) at Bournemouth University, UK.

- mail@eduardzell.com
- www.eduardzell.com

Katja Zibrek

Katja Zibrek is a postdoctoral Research Fellow at Trinity College Dublin (TCD). She holds a diploma in Psychology (University of Ljubljana, Slovenia) and a PhD in Computer Science (TCD). She has conducted research in the area of perception in graphics, particularly in investigation of gender, emotion an personality perception of virtual characters. She has published at SIGGRAPH, ACM Transactions on Applied Perception and IEEE Transactions in Visualisation and Computer Graphics.

- kzibrek@tcd.ie
- www.scss.tcd.ie/~kzibrek/

Rachel McDonnell

Rachel McDonnell is an Assistant Professor at Trinity College Dublin. Her research interests include perception, animation, and virtual humans. She has been a member of many IPCs, including the Eurographics and SIGGRAPH papers committee and has published over 50 papers on topics in perception, facial animation, and virtual humans. She has served as both program and conference chair of the Symposium on Applied Perception, and is on the editorial board for the associated journal - Transactions on Applied Perception.

- ramcdonn@tcd.ie
- https://www.scss.tcd.ie/Rachel.McDonnell/

Xueni Pan

Xueni Pan is an Assistant Professor in VR in Virtual Reality at Goldsmiths College, University of London, and the Programme Lead for MA/MSc in Virtual and Augmented Reality. Over the past 15 years she developed a unique interdisciplinary research profile with journal and conference publications in both VR technology and social neuroscience. Her work has been featured multiple times in the media, including BBC Horizon, the New Scientist magazine, and the Wall Street Journal. She was the co-lead instructor of the Coursera Virtual Reality specialisation consisting of five courses with 25,000 learners internationally, the co-founder and co-chair of the Virtual Social Interaction Workshop.

- x.pan@gold.ac.uk
- http://panxueni.com

Marco Gillies

Dr Marco Gillies is a Reader in Computing and Academic Director: Distance Learning at Goldsmiths University of London. He has over 20 yearsâĂŹ experience of research in Virtual Reality and AI driven Virtual Characters. His research centres on social experiences and embodied interaction for VR and immersive media. He is also a pioneer of the use of AI and machine learning in virtual reality and of the application of Human-Computer Interaction approaches to AI and machine learning. His research is highly interdisciplinary, combining computer science with psychology and the creative arts. This is reflected in research funded by diverse bodies including the Engineering and Physical Science Research Council, the Economics and Social Research Council, the Arts and Humanities Research Council, the Leverhulme Trust, the European Commission, InnovateUK and the Arts Council.

- m.gillies@gold.ac.uk
- https://www.gold.ac.uk/computing/staff/m-gillies/

EXAMPLE COURSE NOTES

4. Visual Perception Basics

Perception is an important part of graphics, and the knowledge of how the human visual system interprets visual stimuli was the foundation of many techniques used in graphics. For example, the knowledge that people can perceive a sequence of images, depicting a moving object, when presented in fast succession as fluid motion, was a foundation for animation. Another example is that people see color due to only three types of light sensitive receptors or "cones" on the retina (red, blue and green). This was the basis of the RGB color system, where any color could be expressed as an integer on the spectrum of red, green and blue. There are many more examples of how graphics exploits these basic traits of the human visual system, however, this course will only make a brief overview of them. The reader is encouraged to refer to the previous courses [FS09, GCL*06b] which cover basic visual perception of 3D environments in more detail. Here, we present a broader definition of perception, which is needed to understand some concepts of virtual character perception.

4.1. Perception

Perception comes from the Latin word perception which literally means "to seize" or "to understand". In order to perceive the world, stimuli need to access the organism through a system, developed to turn information into the activity in the nervous system – a process called "sensation" (visual, tactile, audio, etc.). However, these sensations need to be organized in a meaningful experience. This part is mainly performed by the brain, which processes information from the senses and interprets its relevance to the organism. The first process of transforming sensory–driven information, is also called the **bottom–up** process. The second process is based on acquired information about the world through learning and provides a context for the information from the senses to be interpreted, also known as the **top–down** process. A simple model of the interplay of the processes is shown in Figure 1.

Bottom–up process transforms low–level sensory information into high–level information. This is needed because visual stimuli are very complex. A good example of this process is depth perception – the brain needs to see depth from images, which are displayed in an eye's retina as two dimensional images. So how does the brain do it? Well first, it uses information from two eyes, which deliver a slightly different angle of the image (retinal disparity). Then, it joins the images in the visual cortex to assess the depth information. To do so accurately and fast, it uses a set of learned or predefined rules of organization.

These are usually known as principles of visual organization (also Gestalt principles) which are part of the top-down process. A common principle is "figure and ground" (Rubin [1915], described in [BW58]) where the figure is seen as a meaningful object in the field of view (typically presented as a smaller, connected image), and the ground is the less relevant background. "Grouping" is another form of organization, where our perceptual system joins separate objects together to create a whole by their visual proximity, similarity, continuity, closure, symmetry and common fate [Wer23]. Very important principles regard depth perception. These principles are especially important in art and graphics. For example, to simulate depth in 2D images, (Figure 2a) we can use two lines, converging into a point, which create the illusion that they are actually parallel and continue into a distance. The two yellow lines in the picture have the same length but because they are put on the "path" of the two converging lines they appear to be at different distances, therefore creating the illusion that one is longer than the other. A lot of principles of organization have been observed with these types of visual illusions. The illusions signify that there are competing processes happening in visual perception – for example, in the figure-ground principle, if the visual process cannot determine what is figure and what ground, it will switch between them, depending on the person's attention focus (Figure 2b).

While it is not precisely known which of these principles of visual organization are learned through experience with the world or which are biologically inherited, many still belong to a category of bottom–up perception and are processed in the primary visual cortex, the brain region which processes the most fundamental visual stimuli, such as orientation and color [SSR11]. However, it is known that our top–down processes, such as attention, expectations, motivations, etc., influence our perception as well. These mental representations or "schemas" include everything we learned about the world and provide a fast assessment of the meaning of the stimuli, especially when sensory information is vague or ambiguous. For example, top–down perception is the reason why we can perceive a human form from a simple point-light display [Joh73], due to our previous experience with observing people. However, this ability is also the reason for some erroneous judgments. A practical example is the failure to notice spelling mistakes in the text since we can derive meaning from words even when we put attention on a few letters of those words. On the other hand, these failures can be used to optimize graphics content without introducing a perceptible change (see for example mesh simplifications based on a perceptually driven technique [GH97]).

When perceiving virtual characters, the understanding of both bottom-up and top-down processes are important. Knowing how the 3D shape and depth can be perceived from a 2D representation is just as crucial for character design as understanding the importance of pre-existing schemas when observing virtual humans. These schemas can be very broad and, as we will discuss in the last section of this course, include social perception. In social situations, people frequently make judgments of other people based on very little information. Visual appearance can result in attributing particular personality traits even to strangers, which results in perceiving any behavior from that person in the context of the first judgment. Stereotypes are such an example, where other peoples' behaviour is analyzed according to a group or category they belong to (e.g., race, gender, age) and not unique constellations of their personal attributes. In graphics, we can also use bias as a measure of authenticity of a designed virtual character - a virtual human who is perceived to be realistic in behaviour and appearance, could also induce a biased response [ZKM18].

The recognition of these top–down perceptual effects is also important when designing a perceptual experiment. Participants themselves come from various backgrounds and are in different mental

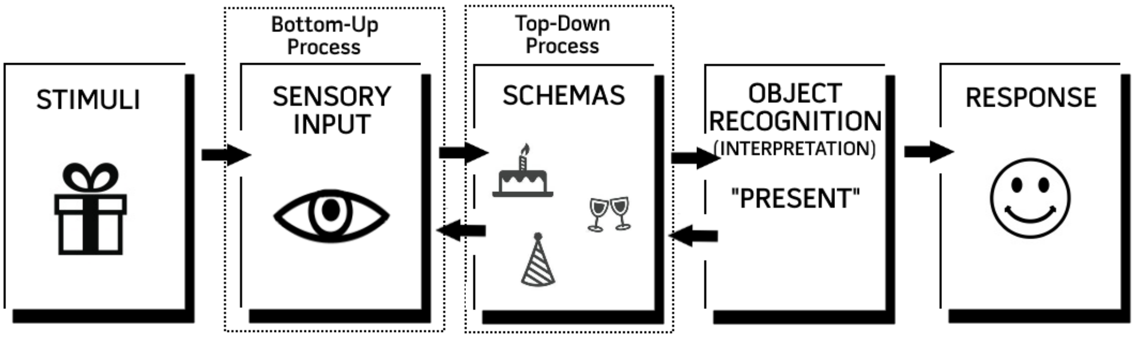

Figure 1: *Simple model of perception. The object is first seen by the sensory system (visual input) and then interpreted in context (being at a birthday party). The object is recognized (present) and ends in the response of happiness. The arrows represent the direction of effect, where there is a feedback loop between sensory information (bottom–up) and already acquired representations of the world (top–down). Top–down processes also include cognitive states of the person (his attention, motivation, personality) which can affect the sensory organ to perceive selectively. For example, if the person does not trust his friends that gave him a present, he might react concerned or even frightened at the sight of the box. There is also an example of perceptual organization (grouping) in this image: the boxes that describe separate steps in the model are registered by the sensory system as broken lines yet we perceive them a rectangular objects, seemingly casting a shadow.*

Figure 2: *Examples of visual illusions: a) Ponzo illusion, where the yellow lines appear to be different even though they are of the same length; b) Rubin vase, where the object in the image can be seen as a vase or two faces.*

states at the time of the experiment, which might interfere with the variables we are measuring. The following section provides a short overview of the basics of experiment design for the investigation of the perception of virtual characters.

4.2. Experiment Design & Statistics

There are some specifics of the experimental design when investigating the perception of virtual characters. In this section, we are going to cover the basics by introducing the following topics:

- The influence of participants (sampling)
- The differences in measures: direct and indirect
- An overview of the most important concepts, when analyzing the data (e.g., mean, standard deviation, trust interval and the difference between significant and intermediate results).
- The importance of controlled stimuli

4.2.1. Participants

As mentioned in the previous section on perception basics, participants come from a variety of backgrounds and this variation can introduce unwanted effects on our collected data. Increasing the number of participants is usually the best way to avoid any effect of individual variation, as this gets dispersed throughout a large sample while only the systematic effects should remain. Another important key is also in the sampling approach - in order to generalize results from the sample of the population, which is the basic premise of inferential statistics used in perceptual studies, the sample should represent the population well. If, for example, we got a result that the recognition of emotion is higher for stylized than realistic characters, and our sample is 30 male and 5 female participants, the conclusion cannot be generalized to all people, only males. It is similar with sampling only from the university campuses, or collecting data from only one cultural group.

In reality, getting a perfect sample for our experiment is usually extremely challenging. However, some improvements can be made simply by insuring that the selection of participants is as randomized as possible. Another way is to include previous knowledge on the perception of people and run pilots or pre-tests to test the experiment design. For example, when investigating recognition of emotion of virtual characters, research on the psychology of emotions shows that the perceivers' own emotions influence emotion perception of others [NHMIK00]. If we are interested to control for this effect, we could measure the emotional state of the participant prior to the experiment and include it in the statistical analysis.

4.2.2. Measures

Peoples' attitudes towards virtual characters can be measured in various different ways. The most commonly used measures are subjective responses, where people are asked to give answers to a questionnaire, such as rate their experience or make a decision about what they had witnessed. Subjective responses are usually obtained by questionnaires, where Likert scales and semantic differential scales are used in the attempt to quantify data. Likert scale prompts the person to give a rating of an agreement with a particular statement (e.g., "On a scale from 1 to 7, how eerie is the character?"), while the semantic differential scale has two different descriptors on each end of the scale, for which it was previously established that they belong to the same dimension (e.g., an emotional response scale can range from happy to sad). These scales have certain disadvantages. People's intention, mood, personality type and other unrelated factors can influence the way they give answers - some people avoid giving extreme ratings, develop ideas and strategies on how to assess stimuli, give intentionally misleading answers, etc. [BCFW08]. In order to control for this, repetitions of the same question can give a more reliable result, and there are even tests which measure people's willingness to give socially desirable answers (Social Desirability Scale, Crowne and Marlowe [CM60]). The most common approach, however, is to use standardized tests, which are created from a set of scales, measuring a specific construct, and have been tested on a large sample and controlled for validity (that the test is measuring the intended construct) and reliability (the test measures the construct consistently across time, individuals and situations). An example of a standardized measure which measures attitudes towards artificial humans is the Godspeed Questionnaire, introduced by Bartneck *et al.* [BKCZ09] and revised by Ho *et al.* [HM10]. This instrument uses 4 indices with high internal reliability - warmth, humanness, eeriness and attractiveness. A lot of research studies, however, do not use the same terminology and the lack of universality is a known issue in the field of character perception [KFMT15].

Another way to avoid subjective mapping of answers, a forced choice task can be used, where a limited range of options is given, and the participant must choose the one which is the closest to his answer. In the Two-Alternative Forced Choice (2AFC) experiment design, speed and accuracy of choices between two alternatives given a timed interval are tested [Bla52]. Most of the low level perceptual experiments use some version of this task, and the goal is to retrieve the thresholds of stimuli detection or the levels of when the stimuli changes the perceived intensity. An example of how this test could be used for evaluating the perception of virtual characters, is in virtual crowds or the so called detection of 'imposters' [HMDO05], where simplified versions of characters are introduced to increase the rendering speed of a large crowd without being noticed by the viewer. An extended version of the 2AFC measure is the multi-dimensional scaling method (MDS), where viewers do not only report the detection of change but also the degree to which the stimuli changes (see for example Logvinenko and Maloney [LM06]).

Brain studies using fMRI and EEG can be used as well. Experimenters can monitor participant's heart rate, respiratory rate and skin conductance to track changes in anxiety levels of people who are observing the character. Peoples' eye gaze can give a lot of information on their attention to particular areas of the character, indicating areas of interest or disturbance. These objective measures are referred to as physiological measures and have many advantages: they are quantifiable and do not require participants' conscious evaluation. Reasons against using these measures could be poor accessibility and cost of the machines, additional expertise for analyzing the results and non-direct association between physiological and mental responses.

Indirect measures are therefore the ones where the participant is not aware of the purpose of the testing and cannot affect the outcome by conscious processing. For example, rather than asking the participant how threatening the character appears to him or her, we can measure participant's increase in heart rate. Other indirect measures are based on semantic priming (e.g., Stroop test [Jen65], Implicit Association test [GMS98]) and are also used to study the perception of virtual characters [BGS13]). The later study also utilizes virtual reality as a tool, where ecological validity of behaviour is possible (e.g., the measure of proximity to virtual humans in the studies of Bailenson *et al.* [BSH*05]). Indirect measures are extremely valuable as they bypass any conscious interpretation from the person which could affect the measured data. However, indirect measures may pose a question to validity - do these measures really reflect the nature of the studied construct? To increase validity, a combination of direct and indirect measures is usually the best choice for a rigorous perceptual study.

Peoples' responses can also be collected through observation and qualitative measures. These methods are helpful when we do not have much knowledge about a particular problem we are investigating and do not know how to approach it. A Q qualitative approach will provide a wide range of data but it will be difficult to analyze in a concise way and subject to noise in the data.

4.3. Stimuli Creation

Obtaining images or videos of adequate quality and that fit the purpose of a perceptual experiment may be a difficult task. In general one is either interested in stimuli that change concisely one single aspect or in a large collection of stimuli, such that inconsistencies will vanish later as variance within the statistical analysis. In the following we list several methods to obtain stimuli of virtual characters together with the advantages and pitfalls of each stimuli creation method.

- Collecting images from the internet may be tempting and is certainly the easiest way to obtain big databases. The downside is that many aspects (backgrounds, light, dresses, image resolution and aspect ratio etc.) cannot be controlled and stimuli taken from blockbuster productions cannot be added to the submission without separate copyright agreements with the publishers, which are difficult to obtain. The more specific the requirements for the stimuli, the more difficult it becomes to find the right images. In some cases consistency can be improved by post-processing the selected images (e.g., replacing the background or color correction).
- Morphing between images or 3d models has the advantage that it can be easily accomplished with specialized software. It is also the fastest method to achieve very fine-grained sampling between a photograph and a virtual character. The downside of this approach is that the interpolation is defined by technical and not artistic terms. Visual artifacts may arise if the images are highly different.
- Creating stimuli manually by a visual artist offers a high level of control on the final result at the cost of being very time-consuming. Caution must be paid to subjective traits that should

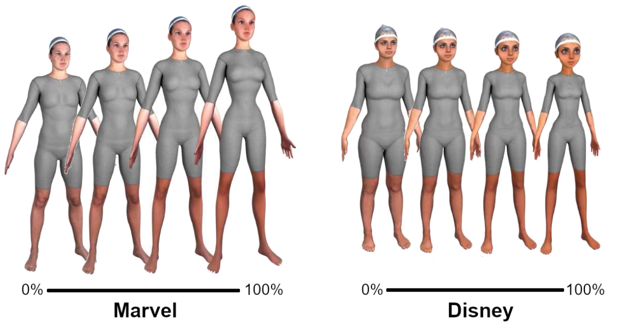

Figure 3: *Transitions between virtual characters from realistic to specific stylizations, sampled by 0%, 33%, 66% and 100% [FMR*16].*

be ideally cross-evaluated to prove that the artist achieved the intended goal.

- Many properties can be controlled in computer graphics by modifying a few parameters. Such parametrizations greatly simplifies creation of highly consistent stimuli (e.g., controlling the light intensity). However, it is unlikely that an equal sampling of the parameter space will also create an equal sampling of the final appearance. The relationship of parameters and the final appearance is in general non-linear.
- For transferring one property (e.g., texture, motion) from one character to another, a mapping or cross-parametrization algorithm will be required. Such algorithms are designed with certain implicit assumptions that may or may not apply to the specific case. For example, algorithms that align 3d meshes try to keep area distortions small. While obvious limitations of the algorithm can be fixed manually, there is also a chance that the algorithm itself introduced unintended effects.

5. Character Appearance

In this section, we will focus on visual characteristics of virtual characters that are largely not affected by temporal changes. Some topics like Level of Detail (Section 5.4) or Lighting and Shading (Section 5.6) have a rich history of general perceptual studies that have only a partial overlap with character related topics. In such cases only a selection of relevant studies will be discussed.

5.1. Character Stylization

Virtual characters are highly diverse, making comparisons between each other a difficult task. Even in terms of categorization, no unique definition exists. Based on the observation of different levels of abstraction in comics, McCloud [McC93] classifies stylization along the iconic and non-iconic scales. A realistic face becomes a smiley under iconic stylization, or a cubist portrait under non-iconic abstraction. Ritchie *et al.* [RCB05] extends this concepts by introducing hyper-realistic characters, a category for characters like The Hulk and Golumn, who look highly realistic but do not exist in real life.

Similar concepts exist for perceptual studies, however a stronger focus is put on stimuli consistency. The two most common scales are photo-realistic vs. stylized/iconic and photo-realistic vs. anthropomorphic or hyper-realistic. It is also common to subdivide the realistic vs. abstract scale further by subdividing the abstraction level into technical components like the general form (shape), resolution of a mesh (tessellation), surface properties (shading, albedo texture etc.) and finally motion. Such subdivision reflects both the different styles in artwork and the technical limitations. Furthermore, the fine grained subdivision allows to track back the contribution and importance of the different ingredients. The downside of this approach is that testing all these parameters becomes difficult, especially because the combinations of parameters grow exponentially. For example testing five characters with four levels of stylization and two different shaders, requires 40 different combinations. If we just increase each scale by one additional sample, the number of stimuli increases to 90. From a practical point of view, creating that many combinations might not even make sense as some combinations would never be used in practice. However, in terms of data analysis, equal sampling is preferred.

Several studies investigated how virtual characters are perceived in terms of realism or appeal across different stylization levels. However, stimuli in early work were mainly based on pictures taken from commercial productions [Han05, Mac06, SWY07, DFH*12]. The lack of consistency (different characters, lighting, backgrounds etc.) of such stimuli, do not always guarantee that the origin of the measured effect is only caused by the different stylizations. For this reason, we focus in the following on studies with more controlled stimuli sets.

Body Perception Fleming *et al.* [FMR*16] evaluated the appeal and realism of female body shapes, which were created as morphs between a realistic character and stylized versions following design principles of major computer animation studios (Figure 3). Surprisingly, the most appealing characters were in-between morphs, where 33% morphs had the highest scores for realism and appeal and 66% morphs were rated as equally appealing, but less realistic.

Faces Wallraven *et al.* [WBC*07] studied the perceived realism, recognition, sincerity, and aesthetics of real and computer-generated facial expressions using 2D filters to provide brush, cartoon, and illustration styles. They concluded that realistic depictions improve subjective certainty about the conveyed expression. Later, they evaluated the perceptual realism of computer-generated faces under progressively blurred normal vectors and textures, finding no effect with their setup [WBCB08]. In the study by McDonnell *et al.* [MBB12] the authors investigated the impact of different rendering styles on the appeal and trustworthiness of the characters (Figure 4, top). In contrast to most studies, this was done for static renderings and short animations. Rendering styles that were close to the most basic shading model in computer graphics were rated as less appealing and trustworthy. Motion amplified this effect. By separating the stylization across shape and material independently, Zell *et al.* [ZAJ*15] (Figure 4, bottom) identified that: (i) Shape is the main descriptor for realism, and material increases realism only in case of realistic shapes. (ii) Strong mismatches in stylization between material and shape negatively affect the appeal and attractiveness of the characters and make them eerier. (iii) The albedo

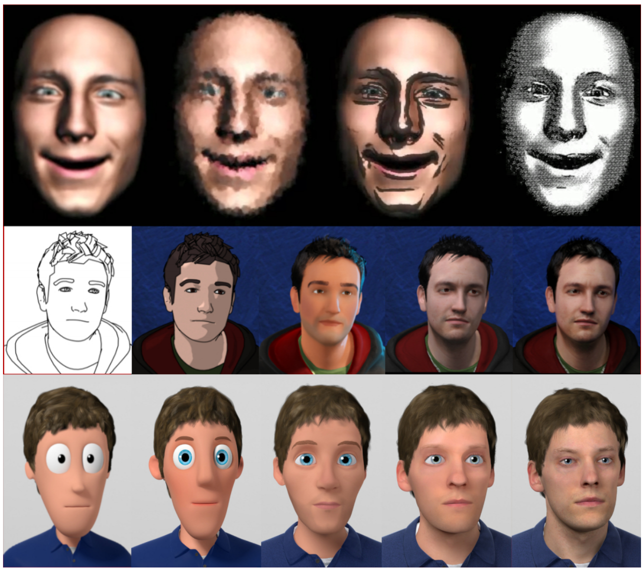

Figure 4: *Examples of consistent stylization.* Top: *non-photorealistic image filtering [WBC*07].* Center: *stylization of rendering styles [MBB12].* Bottom: *stylization of shape and material [ZAJ*15].*

texture modifies primarily the perceived changes of the material and blurring a realistic textures a make-up effect can be achieved - the character appeal and attractiveness increases, without reducing realism. (iv) Finally, abstract characters with realistic materials were perceived as highly eerie, validating the design choices of some horror movies with living puppets.

Instructors One rather frequently encountered use-case of virtual characters are instructors or experts. Despite strong progress over the last years, virtual characters can still be reliably detected in most cases [MBF17]. This raises the question whether they are perceived as competent as real humans. Confronted with an ethical dilemma decision within a medical context, participants had to make a choice before and after the advise of a doctor (expert). The doctor in the video sequence was either a virtual avatar or a human. The recommendation of the doctor had a significant influence on the decision of the participants independently of his appearance. This was even the case when the motion was modified to be "jerky". Similar results were obtained later in a follow-up study [DM18]. However, not all studies come to the same conclusions. Testing the learning outcome of a recorded lecture with slides enriched with a small video of the real instructor, a virtual avatar or a robot, the learning outcome varied. It was smallest for the virtual avatar. Interestingly, students liked the virtual avatar as much as the real person and disliked the robot [LKBJ16]. Another study tested whether stylization level had an influence on expert identification and in consequence whether trust is influenced by stylization [PS16]. Within the study, participants had to answer several difficult question and were assisted by two personalities, where only one was an expert. The personalities were either digital avatars, humans or a humanoid robot. While the robot was placed in front of the wall, the two other were projected on a wall. If the digital avatar was the expert, participants struggled to identify him. In contrast, experts represented by the robot or by a real person were identified reliably. Finally, a meta-analysis comparing the subjectively and objectively measured benefit of adding human-looking virtual avatars as an interface comes to the conclusion that adding an avatar is beneficial, but the effect size is small [YBR07].

Figure 5: *A real person together with a digital double in a virtual environment (CAVE) [WGR*18].*

5.2. Character Realism

With computer graphics reaching closer and closer an indistinguishable level of photo-realism, questions remain on how close we are to this goal and how do we react towards virtual doubles of ourselves.

Becoming Real Earlier studies focused explicitly on the identification of the boundary when characters are perceived as real, by morphing between photographs and puppet faces [LW10] or between photographs and virtual faces [CSJ11]. The results of these works indicate that this question is indeed a categorical decision and that characters must match a high level of realism until they are perceived as real. Furthermore, it seems that eyes and mouth contain the most relevant information followed by the nose, while skin is less relevant (see Section 5.7). Interestingly, due to the visual quality of virtual doubles and output from machine learning algorithms, the identification of computer generated characters is gaining interest within forensic research. Recent studies show that participants exposed to training, feedback within the trial, and incentives were able to classify up to $85-90\%$ of images correctly as real or computer generated compared to a performance of $\approx 50\%$ in mechanical Turk experiments without additional incentives [MBF17]. In terms of error detection, shading is more important than color and compared to other facial areas judging by the eyes alone reveals the highest accuracy [FWN*14].

Doppelgänger Recent developments in 3d scanning facilitate new types of experiments where a realistic virtual double of participants is created during the experiment (Figure 5). Such virtual doubles

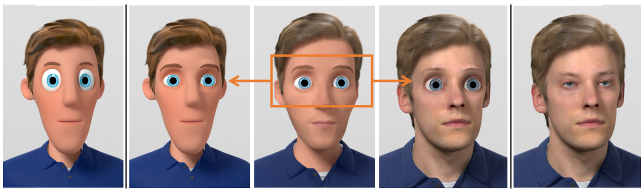

Figure 6: *Illustrative example for changing facial proportions across different stylization levels. Eyes of the character in the middle are transferred to abstract (*left*) and realistic (*right*) characters. While the abstract character still look reasonable, this is not the case for the realistic character.*

or doppelgängers allow testing the importance of having a virtual character that are either a look-alike or non-similar to users. Within an interactive application [FB09], participants could control the weight of a virtual character by doing physical exercise. If the virtual character was a doppelgänger, participants exercised significantly more. Furthermore, advertisement with doppelgängers tends to be more effective [AB11] and participants react less aggressive to doppelgängers of others in games [SB12]. However, watching at a virtual double is not always of benefit. In preparation of public speaking, participants who watched non-similar characters reduced anxiety compared to watching virtual doubles giving the talk [AFKB14]. Virtual doubles, in combination with motion capturing increase body ownership and presence within the virtual environment and facilitate the acceptance of the virtual body as their own [WGR*18, GCHR19]. Finally, accuracy of body weight estimation is independent of the participant's gender [TPS*18]. However, males accepted a larger weight range as their own. Females but not males considered a thinner body as ideal.

5.3. Facial Proportions

In Seyama and Nagayama [SN07] stimuli were created by morphing between photographs of real people and dolls. By controlling individually the morphing speed of facial parts as well as by scaling facial parts individually, it was found that realistic characters were perceived as less appealing if facial parts had strong deviations in terms of size (e.g., eyes have been locally increased – see Figure 6). Several studies confirmed that increasing facial parts lowers perceived appeal, especially in case of realistic characters. In addition, Green *et al.* [GMHV08] demonstrated that not only proportions, but also the placement of facial parts may affect negatively perceived appeal. The measured effect was bigger in cases where the original faces were more attractive and human-like. The results have been later confirmed [MGHK09, BSL13] and partly extended by demonstrating that a mismatch of realism between facial parts negatively affects appeal.

The previous studies addressed the perception of rather unusual facial proportions for realistic characters and their influence on perceived appeal. In case of real people, previous work demonstrated that some facial proportions are associated with personality traits. A meta-analysis study concluded that individuals with wider faces were judged by observers as more threatening, more dominant and less attractive, especially for male faces [GDD*15]. In addition larger eyes increase trustworthiness [ZVC96], while narrow eyes appear aggressively. Parts of these results have been confirmed for virtual characters. Narrow eyes have been rated as more aggressive and less trustworthy for both, abstract creatures [FKM17] and more realistic virtual humans [FM18]. It should be noted that eye size should not be modelled by varying the size of the eyes itself as this will be quickly perceived as eerie and artificial, but rather by changing the shape of the eyelids and partly the proportions of the head. Protruding eyes appear larger, whereas, hooded eyes and monolid eyes appear smaller. For virtual characters, the opposite result was found for the perception of wide faces, which were perceived as less aggressive and dominant [WGH13] even when a masculine rather than a babyface appearance was achieved [FM18].

5.4. Level of Detail

Creating models with different level of detail (LOD) is especially common in real-time applications, crowd scenes and detailed scenes like cities, terrains etc. The overall goal is to maintain rendering speed and a small memory footprint without loosing visual accuracy. Perceptual research on LOD can be divided in two categories: When do humans notice the artifacts of low-quality models and how are low quality models perceived when their smaller resolution is obvious.

Luebke et. al. [LRC*03] provide probably one of the most detailed descriptions on perception within the LOD context and for topics related to crowds we refer to dedicated tutorials [BKA*14, HLLO10, DMTPT09, TOdHCD06, TOY*07]. Within the context of virtual characters, two different representations exist to represent characters at lower resolutions. One option is to use textured, low resolution meshes. The other option are impostor techniques, where the object becomes a plane with albedo textures, transparency and normal maps. To maintain the 3d illusion, textures are rendered from different views and are replaced later depending on the user perspective. For computer displays, impostor representations remain indistinguishable until a pixel-to-texel ratio of slightly above one-to-one [HMDO05]. Flickering artifacts become visible when changing the representation from impostor to 3d mesh at a pixel-to-texel ratio of one-to-one [MDCO06]. Furthermore, impostors are better at reproducing fine scale deformation and subtle motion than low resolution models. Differences in the view direction of $10-20^\circ$ remain unnoticeable. It should be noted that the authors of these studies mentioned visual artifacts due to aliasing as a reason for identification of impostors or transitions. Given the strong improvements in anti-aliasing algorithms, both within the rendered image as well as over the temporal domain, the low quality model may remain longer indistinguishable within current game engines. In the case of low-resolution models, visible artifacts can be identified due to a lack of smoothness within the silhouettes, incorrect lighting and texture distortion. Studying each factor independently for virtual characters, Larkin and O'Sullivan [LO11] showed that silhouette is the dominant artifact for simplification identification at smaller screen spaces and lighting and silhouette artifacts are easily detected at larger screen spaces. However, when using normal maps, lighting artifacts can be masked efficiently.

Perception of low quality models when their smaller resolution is obvious is a side track of realistic vs. non-realistic character percep-

Figure 7: *Appearance change due to facial skin manipulation.* Top: *The original photograph (left) is edited by removing wrinkles (center) and removing aging spots and blurring skin imperfections (left) [FM08].* Bottom: *practical retouching example of the left image with strong editing of the skin appearance (©Rousselos Aravantinos).*

Figure 8: *A special case for demonstrating the ambiguity of light and material. Colored light prevents accurate estimation of the material colors of the porcelain.*

tion research. MacDorman *et al.* [MGHK09] showed participants several images of virtual faces, combining different textures (from realistic to simple lines) with geometric levels of detail, where geometric detail was defined by the polygon count. Results suggested that reducing photo-realism can make the face look less eerie and more attractive. Similarly, Burleigh *et al.* [BSL13] compared faces with enlarged lips and eyes of different mesh resolutions. Faces with the lowest mesh resolution had lower variation on the perceived appeal when comparing normal and increased eye size (see Section 5.3). Furthermore, we also want to point the reader to the general role of perception within the context of mesh compression, where perceptual metrics are increasingly used to control the lossy compression locally [CLL*12].

5.5. Skin Appearance

Besides dedicated work on skin appearance of virtual characters, relevant research has been carried out in the cosmetics research as well as in general research on attractiveness of people. Many studies concerning attractiveness of human faces merged different photographs to achieve average appearance. There was speculation that this technique impacts ratings of attractiveness not just because it averages the shape, but also because it removes blemishes and other skin irregularities [AC91]. Several studies confirmed that texture changes do result in a significantly more attractive face [BP92, LH02]. In the cosmetics domain, Fink *et al.* [FGM06] created textures from photographs of women of different ages and evaluated these textures on a single female virtual character. Renderings with pure skin were rated as younger and more attractive than renderings with strong variations in skin pigmentation. This observation was confirmed in a follow-up study, which showed that blurring the skin texture can increase attractiveness [FM08]. Zell *et al.* [ZAJ*15] observed similar effects for stylized characters as well. Blurring realistic textures, while preserving feature contours (e.g., lip contours) made characters with realistic shape as well as stylized shapes more appealing. In fact, texture stylization can be considered as a process that makes textures more uniform up to the point when features disappear depending on their visual importance (Figure 7). Empirical observations that smoother skin is considered more appealing can also be found in many photograph retouching books (e.g., [NR11]) and photo-retouching software for faces.

5.6. Lighting and Shading

The number of papers addressing different effects of lighting and shading within the context of virtual characters is small, but we can gain valuable knowledge by taking into account general studies on lighting and shading. We focus at this point on glossy, diffuse and translucent surfaces and omit transparent materials as these are less relevant in the context of virtual characters. Studying the perception of materials is challenging due to the strong interaction with lighting conditions. Certain visual appearances can be achieved by either modifying the surface appearance or the environment lighting as demonstrated in the extreme example in (Figure 8). Consider, for example, a perfectly polished chrome ball within a closed box covered with velvet. In the presence of indirect light, the ball would mirror the velvet surface, making it impossible to distinguish it from the velvet material of the box. It is therefore not surprising that participants inconsistently approximate parameters for glossy surfaces, especially in case of unnatural lighting [FDA03]. Also, accuracy of identifying equal materials [PtP06] or the determination of roughness [HLM06] vary for different lighting setups. Besides light, even the shape of an object influences the perception of glossiness [VLD07, Van09, OB10, OB11]. Depending whether a small or big fraction of the surface area is covered by highlights, the material will be perceived as more or less glossy. Such ambiguity can only be resolved by providing several views of the same object. At the same time, the human visual system has developed an incredible ability to account for contextual information as well as surface properties in order to preserve the identity of an observed object. For instance, a black box remains identifiable as black no matter how bright the light within the room is. While humans perceive the box as black under different lighting conditions, the color will range between different shades of dark grey. This adaptation to contextual information is referred to as lightness constancy and it is a major challenge in visual science [Bra03, GKB*99].

In the case of translucent materials, the lighting direction has a fundamental impact on perceived translucency. While frontally lit translucent objects lack many visual cues (e.g., blurred features, soft shadows, low contrast), these features are enhanced when illuminated from the back [FB05]. This effect is exploited in skin and hair rendering, where accurate shading models that replicate the physical behaviour of light are rendered with a back-light to visualize the fidelity of the shading model. In contrast, shaders that focus on performance and sacrifice accuracy are shown under less extreme lighting setups to underline their visual equivalence. The fact that the human visual system is tolerant to inaccuracies in lighting or shading was considered to speed-up rendering, e.g., approximating indirect light between frames through spherical harmonics [JES*12]. Another well-established example is the replacement of the computationally intensive inter-reflections between surfaces through several simpler light sources [Kel97]. Respective perception parameters have been systematically studied in [KFB10].

Rather than focusing on accuracy in material perception, the question remains what makes surfaces and light look realistic. Based on real photographs, Rademacher *et al.* [RLCW01] identified that surface smoothness and shadow softness increase realism, but not the number of objects and lights. However, the effect size was bigger for surface smoothness than for light [Rad02]. A similar approach is considered as good practice among digital artists, who create multi-layered materials, with dedicated textures for surface scratches, dirt etc. It should be noted that it is often sufficient to add plausible dirt textures, but not necessarily replicate exactly the dirt.

Zell *et al.* [ZAJ*15], investigated the perceived differences between accurate lighting and shading models across different stylization levels of virtual characters. Participants judged a character with Phong shading lit by simple directional lights with hard shadows almost as realistic as the same character with complex materials in combination with global illumination. While this result might sound surprising within the computer graphics community, Kardos [Kar34] mentioned 80 years ago that people tend to ignore shadows and shading when describing a scene.

Finally, lighting is considered as a powerful tool in cinematography to emphasize the mood within a scene. So far the majority of conclusions are mainly drawn based on observations [Gro05]. The number of empirical studies testing the conclusions is rather small and does not always align with film theory. For example, a recent empirical study by Poland [Pol15] found that low-key/high contrast stimuli produced lightheartedness, contrary to the beliefs of many theorists and cinematographers. Within the context of virtual characters, Wisessing *et al.* [WDM16] (Figure 9) measured the impact of render style and lighting on the intensity and appeal of expressions in short animation sequences. Different lighting directions, such as the key light placed above or below the character had very little influence on perceived emotional intensity, and dark shadows were rated low on appeal.

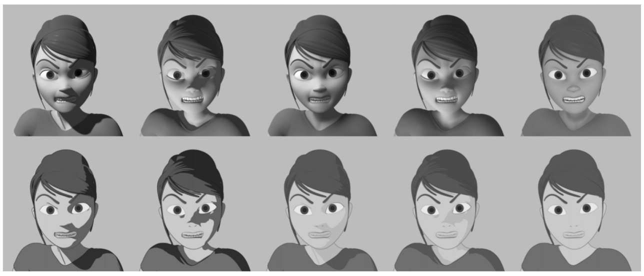

Figure 9: *Stimuli from Wisessing* et al. *[WDM16] for testing two rendering styles (top/bottom) and five lighting conditions. From left to right: High contrast/light from above, High contrast/light from below, Low contrast/light from above, Low contrast/light from below, No directional Light.*

5.7. Visual Attention of Facial and Body parts

Besides focusing on different aspects that contribute towards the appearance of a character, one should also consider that some body parts are more important than others. By using eye-tracking, McDonnell *et al.* [MLH*09] identified that viewers look mainly at the head and upper torso and used this information to create diverse looking crowds more effectively. In the case of faces, it is has been known for a long time [GWG84] that people are looking primarily at the eyes and mouth but the number of fixations at the eyes dominate, which was later confirmed for realistic renderings [RFJ18]. A recent study [SJ16] showed that this is also true for virtual characters of different stylization levels. On average, participants looked for 35% of the time at the eyes, while other regions ranged between 0-10%. This may explain why eyes are considered by practitioners as the most important aspect to achieve realism.

6. Character Motion & Behaviour

Motion or animation of virtual characters is an integral part of the character design and can be achieved through artist animation (keyframing extreme poses and adding in-between frames), physics based animation (computer generated motion based on physics laws), and retrieving the actual motion from real life and applying it to a character (rotoscoping, depth cameras, motion capture), and combinations of those (synthesized motion). Animation approaches depend on the types of use, and they each have perceptually based rules which determine their success in creating a convincing character. In this tutorial, we focus on the perception of realistic motion retrieved from motion capture and retargeted onto a virtual character; we are not interested in all the possible combinations, such as synthesized or procedural motions used in interactive scenarios, since the description of all approaches would result in a lengthy analysis. We therefore leave out some interactive components of virtual characters (see for example approaches to generate artificial eye-gaze behaviour [RPA*15]).

People use non-verbal signs of the character such as motion and appearance to formulate opinions, judgments, or feelings about the character. A lot of social information is expressed through motion. Early work on the perception of biological human motion was done by putting lights on parts of the human body in a darkened room, the so called *point-light displays*. When the human was static, all that was perceived was a group of dots. When moving, the viewers could identify a human body engaged in a readily identified activity, such as walking, running, or dancing [Joh73]. When studying biological motion applied to virtual characters, both shape and motion information interact to formulate a perceptual effect. It was found that the detection of biological motion can be obscured by increasing the anthropomorphism of the character [CHK07].

Motion can also carry information about gender, emotion and personality of the mover. These motion specifics become very important when building virtual characters which are animated using natural motion from motion capture. In the next sections, we expand on these three types of information coming from motion and explain how they affect the perception of characters of different appearance.

6.1. Gender

The research using point-light displays have shown that gender can be recognized from motion when very little shape information is present. Men and women have a specific way of walking, and these differences are apparent to the observers: a pronounced sway in the area of hips often indicates a female walker, while a defined movement in the shoulder area indicates that it is a male walker [KC77]. Not only walking, but also conversational motions (hand gestures, posture) applied to male, female and androgynous characters, can be accurately recognized as male or female motions [MO10]. The participants in this study reported focusing on pose and wrist motions in order to estimate the underlying gender of the mover. Gender can be recognized from facial motions as well [HJ01], where females can be discerned from males primarily because of more frequent nodding, blinking and overall amount of movement [MGCPV07].

Character appearance can affect the perception of gender from motion as well. For motions sparse on gender cues, it is the appearance of the character that will dominate our perception of the character's gender, whereas it is the motion that dominates the perception of characters with an androgynous appearance [MJH*09]. In the case where motion with strong gender cues is applied to a virtual character of the opposite sex, e.g., male walking motion on a female character, it could result in the "contrast effect". Due to this effect, the gender from motion will be perceived even stronger when there is a mismatch with the gender of the character, much like a white paper will appear even whiter when put on a black background. Therefore, a male motion applied to a female character (and vice versa) may actually seem more "manly" due to such contrast [ZHRM15]. Interestingly, this effect was dependent on the gender of the observer - males could identify female motions better and females could identify male motions better on a character of a mismatched sex. This example could point to a selective sensitivity when perceiving gender, perhaps due to the evolutionary importance in correctly recognizing the opposite sex, but also shows the importance of controlling for gender of the participant when conducting perceptual studies.

6.2. Emotion

Because the perception and interpretation of other people's emotion is essential for effective social interaction, people will find the character more engaging when it accurately expresses emotions. And since we put so much importance on emotions in our everyday life, the ability to recognize and distinguish between different emotional states has a prominent role in perceptual processes. Studying the perception of emotion is challenging, as there have been many attempts to define emotions and the exploration of their origin and development is an ongoing research focus [LHJB08]. The most general definition describes emotions as subjective experiences, where the core feeling is that of pleasure or pain [Fri88]. Several approaches to emotion classification exist in the literature as well, from defining emotion as discrete categories [Ekm92] or as dimensions [Meh80, Plu01]. Ekman's approach to emotions was aimed at identifying emotions which are universally recognized and where similarity in their physical expression can be observed. He classified them as basic emotions: anger, happiness, sadness, fear, surprise and disgust. This classification also provided a simple as well as systematical approach for the study of emotion recognition from motion, which provided a comprehensive way to map emotions onto virtual characters. For example, in facial animation, the classification known as the Facial Action Coding System (FACS) is used for the creation of blend-shapes. Dimensional approach to classifying emotions has shown practical use as well, where Russell's circumplex model of emotion [Rus80], describing emotions in terms of valence (positive, negative) and activation (activation, deactivation) led to synthesized motion generation of complex facial expressions for characters [GO06].

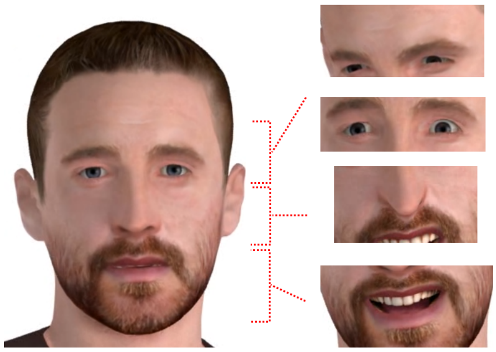

Figure 10: *Areas of the face that carry the most information about emotions: mouth in happy, eyes in fear, disgust in nose and upper mouth areas, while sad and angry are expressed mostly with brow and eyes.*

Basic emotions can be identified through the movement of full body [ADG*04, CG07], upper body [VMD*14] and arm motions [PPBS01], and best emotion recognition rates are achieved

Figure 11: *Emotions were recognized equally well on the different bodies investigated by [MJM*09]: androgynous mannequin, cute character, zombie, and point light display.*

when facial motions are combined [CPM*09, EHEM13, HJO*10], showing that a virtual character will be engaging when both face and body correctly express emotions in the animation. The accuracy of facial expressions is particularly important, as shown by the study of Hodgins *et al.* [HJO*10], where facial motion anomalies were particularly salient even when obvious body motion anomalies were present. Research also shows that there are areas of the face which are important for particular emotions [CKBW04, SPS09]: happiness and surprise are expressed mostly with the mouth; sadness, anger and fear with the eyes and brows, as seen in Figure 10. One particular study integrated this knowledge to reverse-engineer facial expressions and improve recognition of emotions in faces of social robots [CGZ*18]. In terms of body expressions [DM89], the trunk was found to be the most important in conveying positive emotions. In the study of Zibrek *et al.* [ZHRM15], happiness and anger were found to be better expressed with facial and hand movements, while sad and fearful emotions were more apparent in full body motions.

Biological motion can therefore be used to create identifiable emotions for virtual characters. Emotion recognition from body motion is quite robust across different styles of character models, as the study of McDonnell [MJM*09] using a range of virtual characters (Figure 11), from point-light displays to high-fidelity shapes. However, sometimes character's appearance reduces the emotion recognition. For example, more texture information of the skin and some very abstract styles of rendering which introduce a lot of details to the face, interfere with efficient emotion perception from moving faces [WBC*07] or dampen the expression intensity [HCKH13]. While visual fidelity of the character also requires complex texture information, which can dampen the emotion expression, it was found that in behavioral scenarios, it can also change the emotional experience of the viewer. In the study of Volante *et al.* [VBC*16], it was shown that a realistically rendered patient, whose health is slowly deteriorating, unexpectedly increased the feelings of shame and shyness in medical students. The cartoon and sketch rendered styles had a higher value of the expected negative emotions and were more appropriate for inducing a stronger empathetic response, however, the authors suggest that a realistic character added nuance and complexity to the response of the students, making the experience more comparable to a real life situation. A similar finding was reported by Zibrek *et al.* [ZKM18] where viewers were more concerned about the realistically rendered character when he was expressing anger and frustration than other styles (toon CG, toon shaded, zombie, creepy).

6.3. Gender and Emotion bias

When investigating virtual characters, some perceptual effects which are related to our experience in social interactions may arise. Particularly gender stereotypes, which are a consequence of cultural conditioning, impact both the production and perception of emotions [Bre88]. Studies done with subjects living in Western societies show that overall, emotions are perceived to be gender specific. Females were found to be generally more expressive than males (and better at recognizing emotions of others as well) [BPGB05] but certain emotions were more likely to be expressed by males, e.g., anger, contempt and pride [PHKD00]. Sadness and fear were found to be more readily expressed by females [FRMVVM04].

Based on this knowledge from investigating gender differences in perception and generation of emotion, studies investigated whether virtual characters would be subjected to bias as well. The study of Johnson *et al.* [JMP11] explored sex recognition bias on the perception of throwing a ball with an emotional motion style using point-light displays. They found that an angry throw is perceived as more male and a sad throw as more female, which supports the view that anger is more readily attributed to males than females. This view was extended to full body and conversational motions in the study of Zibrek *et al.* [ZHRM15]. Here, gender bias was explored on different types of motion with obvious gender cues (walking) and less obvious cues (conversation), while the motions were applied on male and female virtual characters. They found that emotion biases gender perception according to gender stereotypes: an angry motion is seen as more male, while fear and sadness are seen as less male motions. These studies show that perception of motion is influenced by the type of expression - full body motion has more gender cues than facial or hand motion alone, and emotional expression or appearance of the character will not influence the perception of gender. However, when motion type does not give enough information about the underlying gender, appearance and emotion will both influence the way the biological motion is perceived (see Figure 12).

6.4. Personality

In order to create a more complex and engaging behaviour of virtual characters, one can consider designing them to express personality traits. As with emotion, different personality theories exist in the literature. Due to its continuous examination and re-evaluation, the "Big Five" theory [Gol90, CM92, JNS08] is perceived by many to be the standard description of human personality. The Big Five is a hierarchical model of personality traits with five broad factors (extraversion, agreeableness, conscientiousness, openness to experience, and emotional stability). Each factor is bipolar (e.g., extraversion vs. introversion) and is further described by specific facets and traits. For example, extraverts are talkative and sociable, whereas introverts like to keep to themselves. Emotionally stable

Figure 12: *Examples of emotion expressions from conversational motions (top) and walking motions (bottom) from Zibrek* et al. *[ZHRM15]. Motions from both male and female actors were applied to virtual characters of both sexes to explore the perceptual bias.*

people do not get upset easily and keep a calm attitude even in stressful situations, while neurotic people will be more easily unnerved. People often assess other people on very little information and make judgments about their personality without knowing them before (see studies of zero-acquaintance [BL92, MGP06]): conscientiousness is associated with people who are well groomed and dress smart, while an extraverted person could be identified by having a very expressive face and hand gestures, and just being overall physically animated and energetic. A non-agreeable person could be identified by his or her way of conversing with others - tends to be quarrelsome, callous.

Some of the indicators of particular personality traits are visual (appearance, motion), while others are more associated with language and interactive behaviour. When this information is used to design virtual characters, perception can be steered towards a belief that the character possesses personality traits. It was found that exaggerations in body motion, the general speed of body motion [NWAW10], as well as increased speed of facial motion [HCKH13] create a perception of the character as being more extraverted. A recent study on body shape created by using the skinned multiperson linear model found that for male bodies, extraversion was inferred from their wide shoulders and triangular body shape, while non-agreeable and neurotic personality types in women were associated with heavy-bottom, short-legged powerful body shapes [HPH*18]. Some inadequacies in a character's motion can also transfer to the observed personality disorders of the character. For example, Tinwell [TNC13] found that virtual characters with inadequate upper facial animation exhibit personality traits associated with psychopathy. Gaze behaviour was found to be very closely linked to personality as well [RSSF12, RZM15], where characters who avert eye-contact in a conversation will appear more neurotic/introverted and long, uninterrupted eye-contact a sign of extraversion [RZM15].

It is possible therefore to create distinctive personality traits and these can be systematically investigated. In the study of Zibrek *et al.* [ZM14] for example, the Big Five traits were found to be distinguished by the participants and have a different effect in combination with the render style they were rendered in. In sum, there are many indications that appearance influences the perceived personality, as observed in real people. This creates a fascinating new venue for researchers in character perception to explore.

Figure 13: *Example stimuli from [HRZ*13], where they compared motions of different walkers for distinctiveness.*

6.5. Distinctiveness

Individuals move in different ways, and this distinctiveness in motion is perceived by the observers. Distinctiveness of motion style can be observed in differences between examples of the same behaviour (e.g., slow walk vs. fast walk). When designing virtual humans, human motion style was studied to create motion models, which can be used to add subtle differences to the characters' motion when rendering virtual crowds [MXH*10] or for creating synthetic motion or transitions in motion captured motion [LCR*02]. However, no complete categorisation of motion styles, which would help with the creation of complex motion models, exists yet.

Distinctiveness also helps to identify familiar people based on motion only, which means that motion carries identity cues as well [CK77]. However, this performance is quite poor: in the study of Cutting and Kozlowski [CK77] only 38% of motions were correctly identified, where 16.7% was the chance level of guessing the correct identity. It was also found that people have difficulty distinguishing walking motions of different people [MLD*08] and it was found that only three distinct types of walking motions are necessary to create variety for virtual walkers in a virtual crowd [PO11].

The research on distinctiveness of body motion found a negative relationship with attractiveness - less distinctive movement, created by averaging similar types of motion (walking, jogging, dance) is perceived as more attractive [HRZ*13] (Figure 13). The reason for that is perhaps in the lack of distinctiveness (averageness), which was shown to play an important role in perceived attractiveness of real faces [Rho06]. However, a particular type of distinctiveness, with sexually exaggerated cues, e.g., feminine movement in women characters, has been shown to be attractive as well [JT07].

6.6. From Perception to Interaction

All the studies mentioned in previous sections are related to the perception of non-interactive characters. Since virtual characters appear in various media, they have different levels of interactivity. In computer games, characters need to exhibit some form of intelligent behaviour and responsiveness to the player. In the emerging virtual environments, such as augmented, mixed and virtual reality, believable and interactive characters are central to an immersive experience.

Studies focusing on the perceptual effects of virtual characters in virtual reality have predominantly been exploring the sensation that the character coexists in the space with the user and possesses a form of intelligence [Bio97]. This sensation is referred to as copresence or social presence and is crucial for motivating the user to engage into an interpersonal exchange with the character.

Some studies investigating the effect of character's appearance on social presence found that the anthropomorphic appearance increased social presence [Now01], while others showed no effect of visual realism on social presence [SS02, GSV*03]. A common result suggests that a mismatch between the realism of behaviour and appearance lowers the feeling of social presence [BSH*05] and contributes to a general negative evaluation of the character [GSV*03]. Two of our studies [ZKM17, ZKM18] showed that realistic appearance was in fact preferred in virtual reality, since realistic rendering of the character resulted in higher appeal than stylised renders. No difference in the preferred render style was found when characters were displayed in a non-immersive desktop display. The importance of matching behaviour of the virtual character and appearance of the character was also confirmed [ZKM18]. In this study, the behaviour was presented in a form of personality traits exhibited by the character through verbal and nonverbal cues.

Based on the mentioned research, appearance should be considered when designing virtual characters which create a sensation of social presence. However, the most important component of social presence is interactive behaviour. Bailenson *et al.* [BBBL03] presented a study where a simple eye-contact with the user in virtual reality was enough to create a sense of social presence with the character. More complex behaviour and its effect on the users in virtual reality are the central part of the following sections.

7. Character Interaction

Animated Virtual Characters exist in a range of application areas, from entertainment (films, games) to psychotherapy (for instance, for social anxiety disorders). To date, most commercial applications display Virtual Characters on 2D displays (for instance, Pixar movies), with which users have very limited or no channels to interact with in real-time. The most seen examples of interactive characters are from the gaming industry, where players often engage interactions with computer generated agents (or Non-player Characters, NPCs) as part of the game play, typically through game controllers.

A simple example of such interaction is the classic game "Pacman", developed by Namco (namcofunscape.com) in the 80's, where the player controls Pacman's movements via traditional HCI inputs (keyboard or joystick). The NPCs in the game are the four "Ghosts", who are designed to be chasing the Pacman, whilst displaying different "personalities" (a more in-depth discussion see [Mat03]. However, their behaviour patterns are very limited as the only channel through which the user interact with the NPCs is Pacman's 2D position in the game (controlled by joystick/keyboard). In more recent and more sophisticated games, players could interact with NPCs via other buttons, for instance, by pressing one certain key or a combination of keys the player could fight with monsters, who are programmed to fight back, get "injured", or even "die". More social experiences also emerged in games where players are able to have a "dialogue" with the NPCs by selecting pre-defined answers from a control panel, using keyboard or joysticks. Examples include Mass Effect[†], L.A. Noire[‡], and Heavy Rain[§].

In research, without the limited input constrains from game controllers, a wider range of user inputs can be used to generate real-time character reactions. For instance, the system could capture both verbal and non-verbal cues (gestures, body motion, and gaze) and virtual characters can be programmed to interact with the user in real-time.

Examples in research include:

- The SEMAINE project created four Sensitive Artificial Listeners (SALs) with different personalities (aggressive, cheerful, gloomy, and pragmatic) [Sch11]. All four were autonomous agents and can interact with users in real time, without a human operator.
- The USC Institute for Creative Technologies developed the SimSensei Kiosk with Ellie, an autonomous virtual human interviewer able to engage users for a 20-min interaction where they would feel comfortable to share personal information [DAB*14a]. Ellie is designed to automatically assess user's mental health status and identify issues such as depression, anxiety, or post-traumatic stress disorder.
- The ArticuLab at Carnegie Mellon University has designed social aware robot assistant (SARA), a virtual character who is able to recognize both non-verbal (visual and vocal) and verbal signals and utilizes AI to form her answer [ZSBC16]. SARA's AI is motivated by two goals: task (answering questions, such as help the user to find directions) and social (maintaining a positive and engaging relationship with the user). After a response is formed with its AI model, both verbal and non-verbal behaviours are generated to allow a realistic interaction with the user. All of these systems are focused on generating an emotional connection with the user.

All systems above are typically restricted to a very specific social contexts and can typically discuss pre-trained topics, and the virtual character only able to display a subset of behaviours of a real person. Moreover, similar to applications in gaming, the virtual characters in these systems are also fully autonomous – as in,

† www.ea.com/en-gb/games/mass-effect
‡ www.rockstargames.com/lanoire
§ www.quanticdream.com/en/heavy-rain

their behaviour is entirely driven by computer algorithms. Those autonomous virtual characters are normally defined as *agents* as opposed to *avatars* – which implies real-time representations of a real human, but in the digital format. For instance, in the Pacman game the Pacman is an avatar of the user, while the ghosts are autonomous agents. Often in research there are also semi-autonomous agents – which are virtual characters whose behaviour is partially driven by computer algorithms, partially by another human. The key technical challenge for *avatars* is about representing the user in real-time as much as possible. Autonomous or semi-autonomous *agents* are subjected to a different set of technical challenges – in the following we will focus our discussion on *agents* rather than *avatars*.

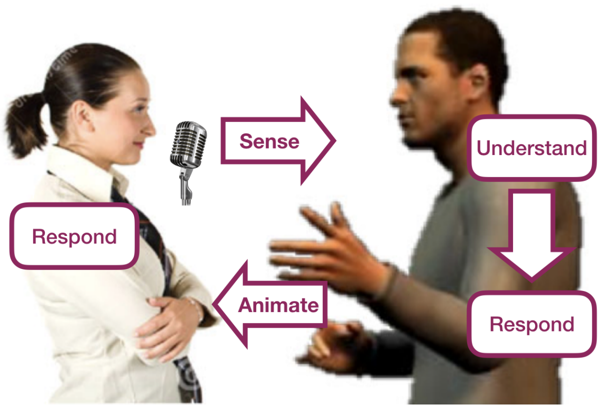

Figure 14: *The human-agent interaction loop*

For an interactive virtual character, it is not enough to do realistic animations or actions, these behaviours must respond to the real person. This involves a loop (Figure 14), the virtual character must sense the the actions of the real human, understand it and produce a response, which in turn will influence the behaviour of the real human. The earlier parts of this tutorial have focused on generating realistic behaviour. What follows will focus on the other aspects of this loop: sensing the behaviour of a human and selecting a response.

Different researchers have used a wide variety of sensing technologies to understand the behaviour of people:

- Video cameras have been very commonly used, with many computer vision algorithms for computing cues such as facial expression or posture.
- Microphones can detect speech: both the verbal content and the non-verbal components such as tone of voice or pitch contours.
- Motion capture or other forms of movement tracking are used to detect body movements, gestures and posture
- Eye trackers can be used to detect gaze direction, and particular eye contact.
- in some cases, physiological sensors can be used to detect emotions that are not obvious from external behaviour.

With the advent of consumer virtual reality, these sensors have become more standardised. Most head mounted displays include head and hand trackers, which can be used to detect gestures and whole body movements. They also commonly include a microphone for speech. Using cameras for facial expression detection is challenging because the head-mounted displays cover the face. Some more advanced systems also include eye trackers and even more exotic sensors such as facial muscle sensors or lip trackers.

All of these sensor data must be interpreted in order to select an appropriate response. Early systems used a "literature-based" approach, in which results from the social psychology or related literature were used to hand craft basic rules for generating behaviour. Many of these rules, though often simple, produced surprisingly effective results. For example, the PIAVCA platform [GPS10] (Figure 20) implemented a number of rules based on the literature. The character would maintain an appropriate social distance and provide feedback signals like nods when spoken to. It would also maintain gaze at the human participant based on a state machine that modelled the conversational turn. The proportion of looking at the participant would vary depending on whether the agent was speaking or listening.

However, it is often difficult to explicitly formulate rules to interpret the complex, nuanced and often subconscious behaviours that humans produce in social interaction. That means that researchers are increasingly turning to machine to understand the behaviour of the human participants in the interaction. For example, Gillies *et al.* [GPSST08, Gil09] developed a pipeline for machine learning of conversational interaction between a human and a virtual agent (Figure 15). The pipeline would begin by recording data from both parties in a conversation, typically motion capture data and audio. The results would be synchronized to create a dataset that contained data of both participants. The result was used to train a machine learning model that mapped from the behaviour of one participant to that of the other. Once trained, the model could be used to generate behaviour of a virtual character based on live input from a human participant.

7.1. Presence and the Illusions of VR

The rise of VR in recent years has been mainly driven by the launch of several lightweight and affordable HMDs from many major consumer electronics companies. All these devices share one common feature: They provide an immersive experience, or "place illusion" [Sla09], defined by (1) 3D stereo vision via two screens – one in front of each eye; (2) surround vision – the real world is "blocked" from your visual perception and as you turn your head you only see the "virtual" world; and (3) user dynamic control of viewpoint which means that the user's head is tracked to update the display in real time according to where the user looks [Bro99]. Implementing these three features together means that the visual information available in VR matches critical properties of the real world, where we have 3D vision all around and the visual scene updates with head movements. These immersive displays which allow us to automatically respond to the computer-generated situation as if they were real, because they support our "sensorimotor contingencies", are commonly described as Immersive Virtual Reality (IVR).

However, these low-level sensorimotor contingencies are not enough to explain the complex interactions that happened between human participants and virtual characters. For that reason, Slater

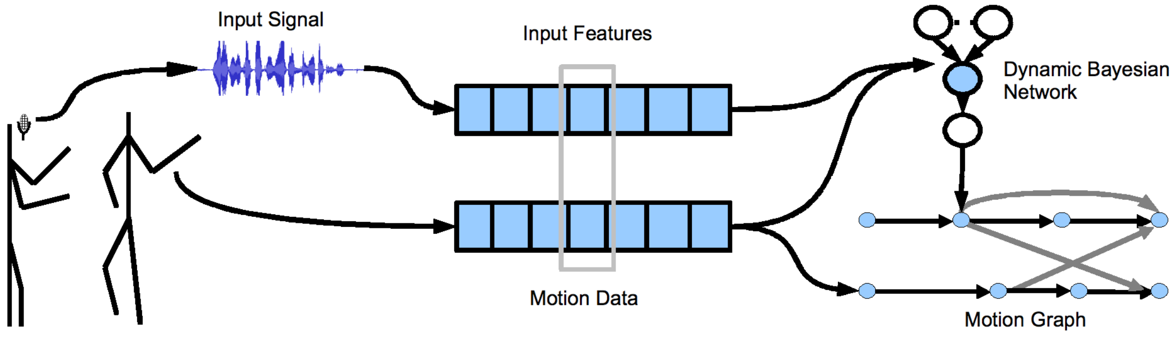

Figure 15: *Two person data capture for machine learning of human-agent interaction. [GPSST08, Gil09]*

introduced a second illusion: *Plausibility*. An interaction with a virtual environment, or in this case a virtual character, is plausible depending on the degree to which if it fulfills three conditions:

- The character responds to the behaviour of the participant
- The behaviour of the character refers personally to the participant
- The character's behaviour is credible in the sense that it fits expectations from similar situations in the real world

Many of the methods described below aim to produce this type of plausible response. For example, in Pan *et al* [PGB*12a] the virtual character fulfils these requirements: her speech and body language both respond to the participant, she speaks directly to the participant and her behavior was scripted to be typical of conversation in a party situation. At the moment, most commercial applications with interactive virtual characters are console or computer games based on a 2D interface (2D display with joysticks or keyboards). However in the following discussions we will focus on interactive characters in immersive experiences enabled by a 3D interface, where the users could express themselves just like they do in real life. This is important in particular for applications in training and therapy, where users are expected to be able to transfer skills learnt from VR to their real life.

7.2. Non-conversational interaction

There are many contexts in which we interact, or simply share space, with other people, without entering into a full conversation. These types of interaction are relatively simple to create in virtual characters, as they do not have the full complexity of conversation, but several studies have shown that they can nonetheless be a powerful experience.

Perhaps the simplest form of this is to share a space without interaction. However, even though we may not interact at all with virtual characters, that does not mean that we do not perceive interaction. For example, Freeman *et al.* [FPA*08] recreated a London underground train carriage (Figure 16) in virtual reality, containing a number of virtual characters, whose behaviour was designed to seem neutral and did not interact with participants. Many participants in their experiment interpreted the characters' behaviour

Figure 16: *Virtual characters on the virtual London Underground train [FPA*08].*

as directed towards them, even though they were not. For example, *"It was nice much nicer than a real experience - people aren't so forthcoming with their feelings in a real situation. Thought they were pretty friendly"*.

A common use of virtual reality is for public speaking training. Pertaub *et al.* [SPBC06] showed that the behaviour of a virtual audience (Figure 17) could have a significant impact on speakers. Participants who gave a talk to a negative audience (characters slouching, looking around, muttering and falling asleep) evaluated their performance much more negatively than those with an audience that smiled and gazed towards the presenter. Anecdotally experimenters noted that participants struggled to talk to the negative audience with some even directly criticizing the audience for their behavior. This shows that even a basic interaction, without true responsiveness, can have a strong impact on people. Recent advances have made it possible to create much more sophisticated and responsive virtual audience, for example, the system by Chollet *et*

Figure 17: *The virtual audience. [SPBC06].*

al. [CWM*15] that can alter it's response based on the performance of the speaker.

7.3. Conversational interaction

While non-conversational interactions can be powerful, the fullest form of human-virtual character can be found in a face to face conversation. Although conversations with on-screen characters and in games are often mediated, for example by typing the statements or selecting options from a dialog tree, the ideal conversation in VR works just as it would in the real world, with verbal interaction (we speak our words directly) and non-verbal (we communicate with body language and tone of voice).

Figure 18: *A male participant interacting with a virtual woman. [PGB*12a].*

A number of studies have shown, that, when possible, this type of interaction can feel very close to a real conversation. For example, male participants having a conversation with a virtual woman (Figure 18) showed initial stress in the early stages of the conversation, as measured by physiological sensors [PGB*12a]. This was particularly pronounced for participants who reported that they were not in a relationship (they were asked this by the virtual character). However, the overall result was a reduction of social anxiety, according to questionnaires administered before and after the experiment. This shows that even a simple conversation with a virtual character can result in real emotional impacts. The behaviour of characters can have an important impact on participants behaviour. Pan *et al.* [PGS15] (Figure 19) asked participants to interview a virtual character, asking potentially embarrassing, personal questions. In one condition, the character had body language and tone of voice that consistent with social anxiety, in the other she appeared very confident. At the end of the interview the character left the room to answer her telephone and did not return. Participants were left to wait alone. They have been told that they could ring a bell to ask her to return. Participants who had interacted with the anxious character waited significantly longer before ringing the bell. This shows that participants can treat virtual characters in accordance with the social rules of politeness and consideration. What is more, this behaviour can be influenced by the behaviour of the character.

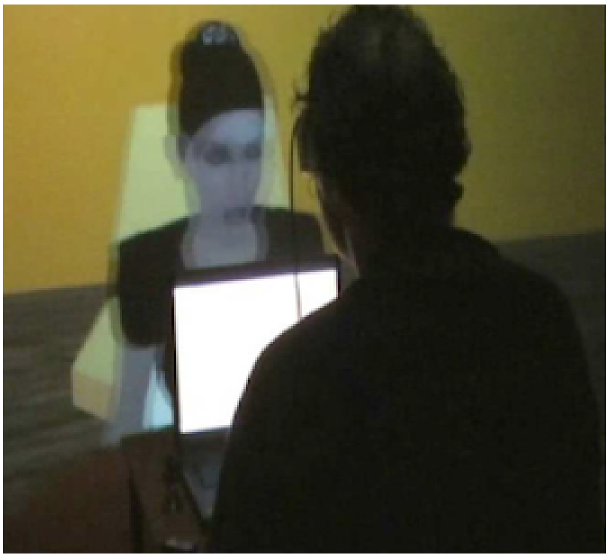

Figure 19: *A participant interviewing a virtual character. [PGS15].*

Human conversation is a highly complex process. It has a sophisticated explicit content, in the form of language, but also a subtle non-verbal subtext in the form of body language, facial expression, gesture, gaze and tone of voice. It involves very tight coordination of behaviour, and in particular rapid turn taking between speaking and listening. To engage full in conversation, a virtual character must interpret speech in real time and formulate responses, while observing the social rules of turn taking and interpreting body language. As described above this involves a tight loop of sensing human behaviour, interpreting it and generating responses. There are many different types of behaviour that make up conversation, and which can influence interaction. The next sections will discuss three aspects that are important and have often been implemented in virtual characters: social mimicry, listening behaviour and verbal interaction.

7.3.1. Mimicry

Capturing participant motion means that the VR environment can be programmed to be responsive in real time, creating the illusion of realistic interactions between the participant and the virtual characters. For instance, knowing the participant's head location means that a virtual character can be programmed to orient their head and/or gaze towards participant's head [FPH16, PH15], and to maintain an appropriate social distance (proxemics) by stepping back or forward [PGB*12b]. The ability to link the behaviour of a virtual character to the participant in real time also facilitated a series of studies on mimicry in VR, where the virtual character copies participants' head movements [BY07, VHPM13], or both head and torso movements [HA16]. Mimicry in these cases are normally implemented with a 1-3 seconds delay in copying. It is also important to note that the psychological effect of mimicry to date is inconclusive. Hale & Hamilton's two studies on mimicry for instance suggested that being mimicked in VR does not always increase rapport or trust [HA16].

7.3.2. Listening

A conversation involves taking turns between speaking and listening. While speaking is the most obvious activity in a conversation, the behaviour of the listener is also vital to maintaining a successful interaction, by providing feedback and encouragement to the speaker, via a range of signals, that are together called the "backchannel". These signals include both verbal (e.g. "yes") and non-verbal utterances (e.g. "uh-huh") as well as other non-verbal signals such as head nods.

A virtual agent that is able to reproduce this type of feedback behaviour is likely to not only be more realistic, but also more likable. For example, Schröder *et al.* [SBC*12] refer to their agents as "Sensitive Artificial Listeners", whose listening behaviour is able to encourage people to speak to them. Similarly, Huang, Morency and Gratch [HMG11] propose that their listening agent is able to build "rapport" with speakers. In both cases, the agents are able to generate effective feedback behaviour without needing to understand the content of speech, by detecting non-verbal features of the speakers behaviour, such as gaze or changes in intonation.

Early systems of this sort were literature-based [GS05,MGM05], using information from the psychology literature to develop simple rules for detecting cues that trigger feedback behaviour such as head nods. These include pauses in speech, head movements or posture shifts. Wang and Gratch [WG10] showed the importance of this kind of feedback. In their study, an agent that looked constantly at the participant while listening was no better than one that ignored the participant, looking around randomly, however, agents that looked while performing appropriate feedback behaviour were rated significantly better than both.

The drawback of this type of literature-based approach, is that the literature only covers those limited aspects of behaviour that generalise best and the behaviours involved are often too complex to be detected with simple rules [HMG11]. This has lead to systems that use machine learning to detect appropriate moments for backchannel behaviour [GPSST08, HMG11, SBC*12]. This requires creating datasets of human listening behaviour. For example, Huang, Morency and Gratch [HMG10] use a process they call Parasocial Consensus Sampling, in which human participants respond to a video clip of a speaker as if they were listening to them live (parasocial interaction). This makes it possible to sample many participants responding to the same clip (consensus). They later showed [HMG11] that a model learned from this data was rated significantly better, by participants interacting with it, than their original literature-based model [MGM05]. The SEMAINE project [SBC*12] used similar techniques to develop listening agents, but rather than focusing on generalizable behaviours, they developed listeners with specific personalities.

7.3.3. Verbal interaction

A full conversation including verbal (language) interaction is one of the greatest challenges for virtual agents. Not only must the agent manage interactions with a human and generate responsive body language, they must also understand the content of human speech and produce appropriate spoken responses. This requires sophisticated natural language processing. Though the language processing techniques themselves are out of the scope of this tutorial, it is worth identifying the challenges involved. For a fully automated conversational interaction, an agent must:

- Recognise human speech from an audio signal and convert it to text
- Interpret the resulting text into a representation that is meaning for the agent
- Formulate a response as text
- Convert the resulting text into speech

Recognising speech has improved dramatically in recent years, and is continuing to do so, though there are still challenges for certain accents and for background noise. However, interpreting that speech is still challenging. For highly constrained domains it is now relatively straightforward, however, the more free-form the conversation, the more challenging it becomes.

From the perspective of generation, once speech has been understood, a response can often be generated with relatively simple rules. However, while converting the resulting text to understandable speech is possible, the generation of realistic and expressive speech is still difficult. While the technology is improving rapidly on all of these tasks, the entire pipeline contains many challenges and most researchers opt to simplify the problem by constraining it in a variety of ways.

The most common response to the problem of realistic speech generation is not to generate speech from scratch but to use a set of audio utterances that have been pre-recording by and actor, thus ensuring realism and expression. Many of the systems described above (e.g. [PGB*12a]) having avoided the problems of natural language understanding by using a wizard-of-oz methodology where a human operator selects utterances for the agent that are appropriate as a response to the participants speech.

However, there are also systems that integrate complex natural language understanding and generation with animated virtual agents, for example, the Virtual Human ToolKit from University of Southern California Institute for Creative Technologies [HTM*13]. This is a very sophisticated system that integrates modules for language understanding, non-verbal signal recognition, body language generation and animation. It has been used in many systems, for example the SimSensei virtual therapist [DAB*14b]. Another example is SARA (the Socially Aware Robot Assistant) [MBZ*16], that is able to use social signals from a human to judge it's current level of rapport with a person and adapt its behaviour.

7.3.4. Integrated conversational behaviour

Conversational must integrate verbal and non-verbal behaviour. For example, gestures produced must be appropriate to speech. For example, SARA [MBZ*16] system just cited is built upon the early BEAT toolkit [CVB01] which is able to generate gesture and other non-verbal behaviours that are appropriate to a spoken utterance.

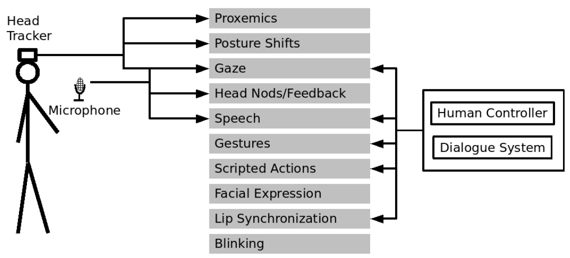

Figure 20: *The PIAVCA model of human-agent interaction. [GPS10].*

Many of these behaviours must act at different time scales and with different levels of reaction to a human participant's behaviour. For example, the PIAVCA platform [GPS10] used in several studies (e.g. [PGB*12a, PGS15] is shown in Figure 20. It includes verbal interaction based using a wizard of oz approach. When an operator selects a particular statement, they are actually triggering a "multimodal utterance" that includes synchronised speech, gesture and other movement (all recorded together from an actress' performance. This verbal interaction proceeds at a relatively slow pace of conversation, with turns changing every few seconds. However, the system also includes automated non-verbal responses which happen much faster. For example, the character responds to speech with head nodding and appropriate eye contact. These behaviours are controlled directly by loop which senses the human's behaviour and generates responses. Some behaviours occur simultaneously with speech, for example posture shifts and gaze continue to respond autonomously during utterances. There is also a turn taking systems, which alternates between pre-recorded utterances and automatically generated listening behaviour. This, relatively simple, integrated conversational architecture shows the complex, parallel multimodal behaviour needed by a conversational agent.

Figure 21: *The child safeguarding. [PCWA*18].*

7.4. Case Studies

7.4.1. Medical Training

Communication training in healthcare is one of the areas interactive virtual humans are often found. For instance, Johnsen *et al.* [JRS*07] showed a significant correlation between medical students' performance with a virtual and a human patient, the study from Raij *et al.* [RJD*07] suggested that medical students were able to elicit the same information from the real and virtual human (although showing less interest and poorer attitude toward the latter). Virtual patients have been used in training for mental health assessment [FCM*15, WBR16], empathetic communication [KRGF*15], and identifying gender bias in diagnosis [RGKK*14]. In these studies, typically, human participants interacted with virtual patients via text or voice, and the virtual patients were animated and programmed to react toward the participants in a realistic way. For a summary of different types of virtual patients see [TSJR12, KZE*15].

We have recently conducted a study to test the effect of using interactive Virtual Characters in an immersive setting for medical training [PCWA*18]. This is a follow up study of our previous work where we tested medical doctors' ability to deny unreasonable demands for antibiotics prescription, using virtual patients, in an Oculus Rift Headset [PSB*16]. As shown in Figure 21, our study [PCWA*18] was conducted in a CAVE-like system where medical doctors could not only see and interact with the virtual patients, but also have access to a physical laptop on which they can read and type notes.

It was a 10 minutes experience, where medical doctors were told to go through a consultation with a virtual patient who had to decide between two sophisticated medical procedures. They were told to read two recommendation letters before the consultation in order to explain to the patient the risks and benefits of each. At the end of the consultation the patient made his decision and left, leaving the medical doctor to type up their notes on the laptop. The patient was a male in his 40s, who came in with his 6 years old son, sitting next to him in the consultation room throughout the process. Unbeknown to the participants, we have animated the father and son to display an abusive relationship: the father exhibited violent behaviour while the child was quiet and withdrawn. We were interested in finding out the extent to which the medical doctors were able to not only notice but also to form a followup plan about the abusive relationship unfolding in front of them. This was evaluated by the quality of the notes they typed up after the consultation (rated by 10 independent raters).

The virtual characters were animated by a combination of motion capture and key-frame animation. They were also programmed to be looking at the participants when listening: in the CAVE-like system participants wear a pair of shuttered glasses with a tracking device on top, so their head position and orientation can be fed into the animation engine in real-time. We pre-recorded over approximate 40 animation clips, each represented as a button in a control panel. A trained experimenter sat in the room and triggered the appropriate response in real-time. This method is often described as "Wizard-of-Oz" as the choice of action appears to come from the virtual character but is actually driven by a human "wizard". The virtual character, in this case, is neither an agent nor an avatar.

We found that medical doctors experiences did not have an impact on their quality of notes. However, their personality did: the quality of their notes is significantly positively correlated to their agreeableness and extroversion, but negatively to neuroticism. It has also been found that the more stressed doctors performed worse than the less stressed ones. These results not only pointed at the value of virtual humans in training, they also highlighted potential strategies to support medical doctors in their dealing with highly sensitive, emotionally charged situations.

7.4.2. VR Games

Virtual reality has now become a mainstream consumer experience, and video games are probably the most popular application of VR in the home. While many VR games simply transfer existing narratives, mechanics and gameplay to a VR setting, VR games hold the possibility of bringing some of the methods discussed in this tutorial to a wide gaming audience, and in the process increase the possibilities for complex social interaction in games. The final case study of this tutorial will discuss work we have done in collaboration with game development studios Maze Theory and Dream Reality Interactive, to develop socially believable VR characters for a commercial game based on the BBC television series "Peaky Blinders".

There are a number of challenges in using the ideas outlined above in a commercial game. The first is that, to be successful, a commercial game must appeal to a wide audience. In part this means that the game should be accessible using a wide range of hardware, which can mean a "lowest common denominator". In 2020, this can actually mean fairly sophisticated VR hardware, with 6 degrees of freedom tracking for head and hands as well as a microphone for voice and other audio input. However, it does exclude other inputs such as face and eye tracking or physiological measures. While speech recognition is possible in principle with audio input, it can possibly exclude many players if it does not work well with certain accents or in a noisy environment.

A second challenge is how the interaction relates to existing game conventions. A game should innovate, but still be recognisable to players. Since game characters do not traditionally have a very realistic social awareness, game players do not normally expect to have to observe traditional social norms, perhaps they will ignore a speaking character and explore the environment. Players need to be guided into taking part in a more realistic social interaction by appropriate story points. Players are also not used to having free form conversations in games. They may not be able to think of appropriate responses and there is a serious risk of them making inappropriate or offensive comments, or otherwise undermine the narrative. This is another reason that speech recognition may not work well.

A third issue is that players must be able to play and complete the game without technical difficulties. As described above recognising non-verbal behaviours in humans is very challenging and can be unreliable. While it might seem appealing to have a scene where a play must appear "charming" to a virtual character in order to win them over, it is likely that recognising "charm" will fail for certain players, either through technical problems or because their individual or cultural way of being "charming" does not correspond to that defined by the game designers. These kind of issues mean that game designers need to be very careful about what kinds of techniques are used as key game mechanics or to make narrative choices. However, it is possible to use less reliable techniques if they do not affect plausibility. For example, game characters can support "micro-interactions" such as nodding when spoken to or making eye contact, which do not affect the narrative but can have a strong impact on the player experience (as shown by many of the studies described above).

To summarise, when making a commercial game it is important to work with the limitations of hardware and the conventions or genre to build robust experiences that work "out of the box" for the vast majority of the population. This means, that at the current state of the art, free form dialogue is unlikely to be effective. While listening behaviour can work more robustly, a character that listens but does not speak is unlikely to drive a narrative forward well. The dialogue needs to primarily be driven by the character, with limited, probably non-verbal, input from the player. The agent should be able to recognise aspects of the player's non-verbal behaviour using the sensors available in a standard VR head mounted display, however, designers must be careful in dividing these interactions into "macro-interactions" that can affect the narrative or gameplay and "micro-interactions". Macro-interactions must be designed carefully so that they are reliable and robust. Players must clearly understand what is expected of them. Micro-interactions on the other hand, can be more subtle and subliminal, and designed to fail gracefully in a way that does not impact the ability to play the game.

Acknowledgements

This research was funded by the Science Foundation Ireland as part of the "Game Face" project (grant 13/CDA/2135) and ADAPT Centre for Digital Content Technology (grant 13/RC/2106) projects. We would like to thank Rousselos Aravantinos, Mario Botsch, Martin Breidt, Bernhard Fink, Thomas Waltemate, Ludovic Hoyet and Christian Wallraven for giving permission to add pictures of their work to this course.

References

[AB11] AHN S. J., BAILENSON J. N.: Self-endorsing versus other-endorsing in virtual environments: The effect on brand preference. *Jour-*

nal of Advertising (2011), 10–2753. 9

[AC91] ALLEY T. R., CUNNINGHAM M. R.: Averaged faces are attractive, but very attractive faces are not average. *Psychological Science 2*, 2 (1991), 123–125. 10

[ADG*04] ATKINSON A. P., DITTRICH W. H., GEMMELL A. J., YOUNG A. W., ET AL.: Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception 33*, 6 (2004), 717–746. 12

[AFKB14] AYMERICH-FRANCH L., KIZILCEC R. F., BAILENSON J. N.: The relationship between virtual self similarity and social anxiety. *Frontiers in Human Neuroscience 8* (2014), 944. doi:10.3389/fnhum.2014.00944. 9

[BBBL03] BAILENSON J. N., BLASCOVICH J., BEALL A. C., LOOMIS J. M.: Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin 29*, 7 (2003), 819–833. 15

[BCFW08] BARTZ D., CUNNINGHAM D., FISCHER J., WALLRAVEN C.: The role of perception for computer graphics. *Eurographics state-of-the-art-reports* (2008), 65–86. 6

[BGS13] BANAKOU D., GROTEN R., SLATER M.: Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proc. of the National Academy of Sciences 110*, 31 (2013), 12846–12851. 6

[Bio97] BIOCCA F.: The cyborg's dilemma: Progressive embodiment in virtual environments [1]. *Journal of Computer-Mediated Communication 3*, 2 (1997), 0–0. 15

[BKA*14] BADLER N., KAPADIA M., ALLBECK J. M., CHRYSANTHOU Y. L., PELECHANO N., GUY S.: Simulating heterogeneous crowds with interactive behaviors. In *Eurographics Tutorials* (2014). 2, 9

[BKCZ09] BARTNECK C., KULIĆ D., CROFT E., ZOGHBI S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics 1*, 1 (2009), 71–81. 6

[BL92] BORKENAU P., LIEBLER A.: Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology 62*, 4 (1992), 645. 14

[Bla52] BLACKWELL H. R.: Studies of psychophysical methods for measuring visual thresholds. *JOSA 42*, 9 (1952), 606–614. 6

[BP92] BENSON P., PERRETT D.: Face to face with the perfect image. *New Scientist 133* (1992), 1809. 10

[BPGB05] BATTOCCHI A., PIANESI F., GOREN-BAR D.: A first evaluation study of a database of kinetic facial expressions (DaFEx). In *Proc. of Int. Conf. on Multimodal Interfaces* (2005), ICMI '05, pp. 214–221. 13

[Bra03] BRAINARD D. H.: Color appearance and color difference specification. *The Science of Color*, 2 (2003), 191–216. 11

[Bre88] BREWER M. B.: *A dual process model of impression formation*. Lawrence Erlbaum Associates, Inc, 1988. 13

[Bro99] BROOKS F. P.: What's real about virtual reality? *IEEE Computer graphics and applications 19*, 6 (1999), 16–27. 16

[BSH*05] BAILENSON J. N., SWINTH K. R., HOYT C. L., PERSKY S., DIMOV A., BLASCOVICH J.: The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence 14*, 4 (2005), 379–393. 6, 15

[BSL13] BURLEIGH T. J., SCHOENHERR J. R., LACROIX G. L.: Does the Uncanny Valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in Human Behavior 29*, 3 (2013), 759–771. 9, 10

[BW58] BEARDSLEE D. C., WERTHEIMER M. E.: *Readings in perception.* D. Van Nostrand, 1958. 4

[BY07] BAILENSON J. N., YEE N.: Virtual interpersonal touch and digital chameleons. *Journal of Nonverbal Behavior 31*, 4 (2007), 225–242. 19

[CG07] CRANE E., GROSS M.: Motion capture and emotion: Affect detection in whole body movement. In *Affective computing and intelligent interaction*. Springer, 2007, pp. 95–101. 12

[CGZ*18] CHEN C., GARROD O. G., ZHAN J., BESKOW J., SCHYNS P. G., JACK R. E.: Reverse engineering psychologically valid facial expressions of emotion into social robots. In *Int. Conf. on Automatic Face & Gesture Recognition* (2018), pp. 448–452. 13

[CHK07] CHAMINADE T., HODGINS J., KAWATO M.: Anthropomorphism influences perception of computer-animated characters' actions. *Social Cognitive and Affective Neuroscience 2*, 3 (2007). 12

[CK77] CUTTING J. E., KOZLOWSKI L. T.: Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the psychonomic society 9*, 5 (1977), 353–356. 14

[CKBW04] CUNNINGHAM D. W., KLEINER M., BÜLTHOFF H. H., WALLRAVEN C.: The components of conversational facial expressions. In *Proc. of Symp. on Applied Perception in Graphics and Visualization* (2004), ACM, pp. 143–150. 13

[CLL*12] CORSINI M., LARABI M.-C., LAVOUÉ G., PETRÍK O., VÁSA L., WANG K.: Perceptual metrics for static and dynamic triangle meshes. In *Eurographics STAR* (2012), pp. 135–157. 10

[CM60] CROWNE D. P., MARLOWE D.: A new scale of social desirability independent of psychopathology. *Journal of consulting psychology 24*, 4 (1960), 349. 6

[CM92] COSTA P. T., MCCRAE R. R.: Professional manual: revised neo personality inventory (NEO-PI-R) and neo five-factor inventory (NEO-FFI). *Odessa, FL: Psychological Assessment Resources* (1992). 13

[CPM*09] CLAVEL C., PLESSIER J., MARTIN J.-C., ACH L., MOREL B.: Combining facial and postural expressions of emotions in a virtual character. In *Proc. of Int. Conf. on Intelligent Virtual Agents* (2009), IVA '09, pp. 287–300. 13

[CSJ11] CHEETHAM M., SUTER P., JÄNCKE L.: The human likeness dimension of the Uncanny Valley hypothesis: Behavioral and functional MRI findings. *Frontiers in Human Neuroscience 5* (2011), 126. 8

[CVB01] CASSELL J., VILHJÁLMSSON H. H., BICKMORE T.: BEAT: The Behavior Expression Animation Toolkit. In *Proceedings of ACM SIGGRAPH 2001* (aug 2001), Computer Graphics Proceedings, Annual Conference Series, pp. 477–486. 20

[CW13] CUNNINGHAM D., WALLRAVEN C.: Understanding and designing perceptual experiments. In *Eurographics Tutorial* (2013). 2

[CWM*15] CHOLLET M., WÖRTWEIN T., MORENCY L. P., SHAPIRO A., SCHERER S.: Exploring feedback strategies to improve public speaking: An interactive virtual audience framework. *UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, November (2015), 1143–1154. doi:10.1145/2750858.2806060. 18

[DAB*14a] DEVAULT D., ARTSTEIN R., BENN G., DEY T., FAST E., GAINER A., GEORGILA K., GRATCH J., HARTHOLT A., LHOMMET M., ET AL.: Simsensei kiosk: A virtual human interviewer for healthcare decision support. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems* (2014), pp. 1061–1068. 15

[DAB*14b] DEVAULT D., ARTSTEIN R., BENN G., DEY T., FAST E., GAINER A., GEORGILA K., GRATCH J., HARTHOLT A., LHOMMET M., LUCAS G., MARSELLA S., MORBINI F., NAZARIAN A., SCHERER S., STRATOU G., SURI A., TRAUM D., WOOD R., XU Y., RIZZO A., MORENCY L. P.: SimSensei kiosk: A virtual human interviewer for healthcare decision support. In *13th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2014* (2014), vol. 2, pp. 1061–1068. URL: http://simsensei.ict.usc.edu/. 20

[DFH*12] DILL V., FLACH L. M., HOCEVAR R., LYKAWKA C., MUSSE S. R., PINHO M. S.: Evaluation of the Uncanny Valley in CG characters. In *Proc. of Intelligent Virtual Agents* (2012), pp. 511–513. 7

[DM89] De Meijer M.: The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal behavior 13*, 4 (1989), 247–268. 13

[DM18] Dai Z., MacDorman K. F.: The doctor's digital double: how warmth, competence, and animation promote adherence intention. *PeerJ Computer Science 4* (2018), e168. 8

[DMTPT09] Donikian S., Magnenat-Thalmann N., Pettré J., Thalmann D.: Course: Modeling individualities in groups and crowds. In *Eurographics Tutorials* (2009). 2, 9

[EHEM13] Ennis C., Hoyet L., Egges A., McDonnell R.: Emotion capture: Emotionally expressive characters for games. In *Proc. of Motion on Games* (2013), ACM, pp. 53–60. 13

[Ekm92] Ekman P.: An argument for basic emotions. *Cognition and Emotion 6*, 3-4 (1992), 169–200. 12

[FB05] Fleming R. W., Bülthoff H. H.: Low-level image cues in the perception of translucent materials. *ACM Transactions on Applied Perception 2*, 3 (2005), 346–382. 11

[FB09] Fox J., Bailenson J. N.: Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors. *Media Psychology 12* (2009), 1–25. 9

[FCM*15] Foster A., Chaudhary N., Murphy J., Lok B., Waller J., Buckley P. F.: The use of simulation to teach suicide risk assessment to health profession traineesâĂŤrationale, methodology, and a proof of concept demonstration with a virtual patient. *Academic psychiatry 39*, 6 (2015), 620–629. 20

[FDA03] Fleming R. W., Dror R. O., Adelson E. H.: Real-world illumination and the perception of surface reflectance properties. *Journal of Vision 3*, 6 (2003), 347–368. 10

[FGM06] Fink B., Grammer K., Matts P. J.: Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evolution and Human Behavior 27*, 6 (2006), 845–857. 10

[FKM17] Ferstl Y., Kokkinara E., McDonnell R.: Facial features of non-player creatures can influence moral decisions in video games. *ACM Transaction on Applied Perception 15*, 1 (2017), 4:1–4:12. doi:10.1145/3129561. 9

[FM08] Fink B., Matts P. J.: The effects of skin colour distribution and topography cues on the perception of female facial age and health. *Journal of the European Academy of Dermatology and Venereology 22*, 4 (2008), 493–498. 10

[FM18] Ferstl Y., McDonnell R.: A perceptual study on the manipulation of facial features for trait portrayal in virtual agents. In *Proc. of Int. Conf. on Intelligent Virtual Agents (IVA)* (2018), pp. 281–288. doi:10.1145/3267851.3267891. 9

[FMR*16] Fleming R., Mohler B. J., Romero J., Black M. J., Breidt M.: Appealing female avatars from 3D body scans: Perceptual effects of stylization. In *Int. Conf. on Computer Graphics Theory and Applications (GRAPP)* (2016). 7

[FPA*08] Freeman D., Pugh K., Antley A., Slater M., Bebbington P., Gittins M., Dunn G., Kuipers E., Fowler D., Garety P.: Virtual reality study of paranoid thinking in the general population. *British Journal of Psychiatry 192*, 4 (2008), 258–263. doi:10.1192/bjp.bp.107.044677. 17

[FPH16] Forbes P. A., Pan X., Hamilton A. F. d. C.: Reduced mimicry to virtual reality avatars in autism spectrum disorder. *Journal of autism and developmental disorders 46*, 12 (2016), 3788–3797. 19

[Fri88] Frijda N. H.: The laws of emotion. *American psychologist 43*, 5 (1988), 349. 12

[FRMVVM04] Fischer A. H., Rodriguez Mosquera P. M., Van Vianen A. E., Manstead A. S.: Gender and culture differences in emotion. *Emotion 4*, 1 (2004), 87. 13

[FS09] Fleming R. W., Singh M.: Visual perception of 3d shape. In *ACM SIGGRAPH Courses* (2009), pp. 24:1–24:94. 2, 4

[FWN*14] Fan S., Wang R., Ng T.-T., Tan C. Y.-C., Herberg J. S., Koenig B. L.: Human perception of visual realism for photo and computer-generated face images. *ACM Transaction Applied Perception 11*, 2 (July 2014), 7:1–7:21. doi:10.1145/2620030. 8

[GCHR19] Gorisse G., Christmann O., Houzangbe S., Richir S.: From robot to virtual doppelganger: Impact of visual fidelity of avatars controlled in third-person perspective on embodiment and behavior in immersive virtual environments. *Frontiers in Robotics and AI 6* (2019), 8. doi:10.3389/frobt.2019.00008. 9

[GCL*06a] Glencross M., Chalmers A. G., Lin M. C., Lin M. C., Otaduy M. A., Gutierrez D.: Exploiting perception in high-fidelity virtual environments. In *ACM SIGGRAPH 2006 Courses* (2006). 2

[GCL*06b] Glencross M., Chalmers A. G., Lin M. C., Lin M. C., Otaduy M. A., Gutierrez D.: Exploiting perception in high-fidelity virtual environments. In *ACM SIGGRAPH 2006 Courses* (2006), SIGGRAPH '06. URL: http://doi.acm.org/10.1145/1185657.1185814. 4

[GDD*15] Geniole S. N., Denson T. F., Dixson B. J., Carré J. M., McCormick C. M.: Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. *PloS one 10*, 7 (2015), e0132726. 9

[Ges12] Geslin E.: Method of induction of basic and complex emotions in video games and virtual environments. In *SIGGRAPH Asia 2012 Courses* (2012), pp. 6:1–6:70. 2

[GH97] Gibson S., Hubbold R. J.: Perceptually-driven radiosity. *Computer Graphics Forum 16*, 2 (1997), 129–141. 4

[Gil09] Gillies M.: Learning Finite-State Machine Controllers From Motion Capture Data. *IEEE Transactions of Computational Intelligence and AI in Games 1* (2009), 63–72. 16, 17

[GKB*99] Gilchrist A., Kossyfidis C., Bonato F., Agostini T., Cataliotti J., Li X., Spehar B., Annan V., Economou E.: An anchoring theory of lightness perception. *Psychological Review 106*, 4 (1999). 11

[GMHV08] Green R. D., MacDorman K. F., Ho C.-C., Vasudevan S.: Sensitivity to the proportions of faces that vary in human likeness. *Computers in Human Behavior 24*, 5 (2008), 2456–2474. 9

[GMS98] Greenwald A. G., McGhee D. E., Schwartz J. L.: Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology 74*, 6 (1998), 1464. 6

[GO06] Grammer K., Oberzaucher E.: The reconstruction of facial expressions in embodied systems. *ZiF – Mitteilungen 2* (2006), 14–31. 12

[Gol90] Goldberg L. R.: An alternative "description of personality": the big-five factor structure. *Journal of personality and social psychology 59*, 6 (1990), 1216. 13

[GPS10] Gillies M., Pan X., Slater M.: Piavca: a framework for heterogeneous interactions with virtual characters. *Virtual Reality* (2010), 1–8. 16, 20

[GPSST08] Gillies M., Pan X., Slater M., Shawe-Taylor J.: Responsive Listening Behavior. *Computer Animation and Virtual Worlds 19* (2008), 1–11. 16, 17, 19

[Gro05] Grodal T.: *Film Lighting and Mood*. Southern Illinois University Press, United States, 2005, pp. 152–163. 11

[GS05] Gillies M., Slater M.: Non-verbal Communication for Correlational Characters. In *Proceedings of the 8th Annual International Workshop on Presence* (2005). 19

[GSV*03] Garau M., Slater M., Vinayagamoorthy V., Brogni A., Steed A., Sasse M. A.: The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2003), ACM, pp. 529–536. 15

[GWG84] GRONER R., WALDER F., GRONER M.: Looking at faces: local and global aspects of scanpaths. *Theoretical and applied aspects of eye movements research* (1984), 523–533. 11

[HA16] HALE J., ANTONIA F. D. C.: Testing the relationship between mimicry, trust and rapport in virtual reality conversations. *Scientific reports 6* (2016), 35295. 19

[Han05] HANSON D.: Expanding the aesthetics possibilities for humanlike robots. In *Proc. of IEEE Humanoid Robotics Conf., Special Session on the Uncanny Valley* (2005). 7

[HCKH13] HYDE J., CARTER E. J., KIESLER S., HODGINS J. K.: Perceptual effects of damped and exaggerated facial motion in animated characters. In *Proc. of IEEE Int. Conf. on Automatic Face and Gesture Recognition* (2013). 13, 14

[HJ01] HILL H., JOHNSTON A.: Categorizing sex and identity from the biological motion of faces. *Current biology 11*, 11 (2001), 880–885. 12

[HJO*10] HODGINS J., JÖRG S., O'SULLIVAN C., PARK S. I., MAHLER M.: The saliency of anomalies in animated human characters. *ACM Transactions on Applied Perception 7*, 4 (2010). 13

[HLLO10] HUERRE S., LEE J., LIN M., O'SULLIVAN C.: Simulating believable crowd and group behaviors. In *ACM SIGGRAPH ASIA 2010 Courses* (2010), pp. 13:1–13:92. 2, 9

[HLM06] HO Y.-X., LANDY M. S., MALONEY L. T.: How direction of illumination affects visually perceived surface roughness. *Journal of Vision 6*, 8 (2006), 634–648. 10

[HM10] HO C.-C., MACDORMAN K. F.: Revisiting the Uncanny Valley theory: Developing and validating an alternative to the godspeed indices. *Computers in Human Behavior 26*, 6 (2010), 1508–1518. 6

[HMDO05] HAMILL J., MCDONNELL R., DOBBYN S., O'SULLIVAN C.: Perceptual evaluation of impostor representations for virtual humans and buildings. *Computer Graphics Forum 24*, 3 (2005), 623–633. 6, 9

[HMG10] HUANG L., MORENCY L.-P., GRATCH J.: Learning backchannel prediction model from parasocial consensus sampling: a subjective evaluation. In *Proceedings of the 10th international conference on Intelligent virtual agents* (Berlin, Heidelberg, 2010), IVA'10, Springer-Verlag, pp. 159–172. 19

[HMG11] HUANG L., MORENCY L. P., GRATCH J.: Virtual rapport 2.0. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2011), vol. 6895 LNAI, pp. 68–79. doi:10.1007/978-3-642-23974-8_8. 19

[HOP09] HERTZMANN A., O'SULLIVAN C., PERLIN K.: Realistic human body movement for emotional expressiveness. In *ACM SIGGRAPH 2009 Courses* (2009), pp. 1–27. 2

[HPH*18] HU Y., PARDE C. J., HILL M. Q., MAHMOOD N., O'TOOLE A. J.: First impressions of personality traits from body shapes. *Psychological Science 29*, 12 (2018), 1969–1983. 14

[HRZ*13] HOYET L., RYALL K., ZIBREK K., PARK H., LEE J., HODGINS J., O'SULLIVAN C.: Evaluating the distinctiveness and attractiveness of human motions on realistic virtual bodies. *ACM Transaction Graphics 32*, 6 (2013), 204:1–204:11. 14

[HTM*13] HARTHOLT A., TRAUM D., MARSELLA S. C., SHAPIRO A., STRATOU G., LEUSKI A., MORENCY L.-P., GRATCH J.: All Together Now: Introducing the Virtual Human Toolkit. In *13th International Conference on Intelligent Virtual Agents* (Edinburgh, UK, Aug. 2013). 20

[Jen65] JENSEN A. R.: Scoring the stroop test. *Acta psychologica 24*, 5 (1965), 398–408. 6

[JES*12] JARABO A., EYCK T. V., SUNDSTEDT V., BALA K., GUTIERREZ D., O'SULLIVAN C.: Crowd light: Evaluating the perceived fidelity of illuminated dynamic scenes. *Computer Graphics Forum 31*, 2 (2012), 565–574. 11

[JKF*11] JUNG Y., KUIJPER A., FELLNER D. W., KIPP M., MIKSATKO J., GRATCH J., THALMANN D.: Believable virtual characters in human-computer dialogs. In *Eurographics STAR* (2011). 2

[JMP11] JOHNSON K. L., MCKAY L. S., POLLICK F. E.: He throws like a girl (but only when he's sad): Emotion affects sex-decoding of biological motion displays. *Cognition 119*, 2 (2011), 265–280. 13

[JNS08] JOHN O. P., NAUMANN L. P., SOTO C. J.: Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research 3* (2008), 114–158. 13

[Joh73] JOHANSSON G.: Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics 14*, 2 (1973), 201–211. 4, 12

[JRS*07] JOHNSEN K., RAIJ A., STEVENS A., LIND D. S., LOK B.: The validity of a virtual human experience for interpersonal skills education. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2007), pp. 1049–1058. 20

[JT07] JOHNSON K. L., TASSINARY L. G.: Compatibility of basic social perceptions determines perceived attractiveness. *Proc. of the National Academy of Sciences 104*, 12 (2007), 5246–5251. 15

[Kar34] KARDOS L.: Ding und schatten. eine experimentelle untersuchung Ãijber die grundlagen des farbensehens. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* (1934). 11

[KC77] KOZLOWSKI L. T., CUTTING J. E.: Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics 21*, 6 (1977), 575–580. 12

[Kel97] KELLER A.: Instant radiosity. In *Proc. of ACM SIGGRAPH* (1997), pp. 49–56. 11

[KFB10] KŘIVÁNEK J., FERWERDA J. A., BALA K.: Effects of global illumination approximations on material appearance. *ACM Transaction on Graphics 29*, 4 (2010), 112:1–112:10. 11

[KFMT15] KÄTSYRI J., FÖRGER K., MÄKÄRÄINEN M., TAKALA T.: A review of empirical evidence on different Uncanny Valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology 6* (2015), 390. 6

[KRGF*15] KLEINSMITH A., RIVERA-GUTIERREZ D., FINNEY G., CENDAN J., LOK B.: Understanding empathy training with virtual patients. *Computers in human behavior 52* (2015), 151–158. 20

[KZE*15] KONONOWICZ A. A., ZARY N., EDELBRING S., CORRAL J., HEGE I.: Virtual patients-what are we talking about? a framework to classify the meanings of the term in healthcare education. *BMC medical education 15*, 1 (2015), 11. 20

[LCR*02] LEE J., CHAI J., REITSMA P. S., HODGINS J. K., POLLARD N. S.: Interactive control of avatars animated with human motion data. *ACM Transaction on Graphics (TOG) 21*, 3 (2002), 491–500. 14

[LH02] LITTLE A. C., HANCOCK P. J.: The role of masculinity and distinctiveness in judgements of human male facial attractiveness. *British Journal of Psychology 93*, 4 (2002), 451–464. 10

[LHJB08] LEWIS M., HAVILAND-JONES J. M., BARRETT L. F.: *Handbook of Emotions, Third Edition*. Guilford Publications, 2008. 12

[LKBJ16] LI J., KIZILCEC R., BAILENSON J., JU W.: Social robots and virtual agents as lecturers for video instruction. *Computers in Human Behavior 55* (2016), 1222 – 1230. doi:https://doi.org/10.1016/j.chb.2015.04.005. 8

[LM06] LOGVINENKO A. D., MALONEY L. T.: The proximity structure of achromatic surface colors and the impossibility of asymmetric lightness matching. *Perception & Psychophysics 68*, 1 (2006), 76–83. 6

[LO11] LARKIN M., O'SULLIVAN C.: Perception of simplification artifacts for animated characters. In *Proc. of Symp. on Applied Perception in Graphics and Visualization* (2011), APGV '11, pp. 93–100. doi:10.1145/2077451.2077469. 9

[LRC*03] LUEBKE D., REDDY M., COHEN J. D., VARSHNEY A., WATSON B., HUEBNER R.: *Level of Detail for 3D Graphics*. Morgan Kaufmann, 2003. 9

[LW10] LOOSER C. E., WHEATLEY T.: The tipping point of animacy how, when, and where we perceive life in a face. *Psychological Science 21*, 12 (2010), 1854–1862. 8

[Mac06] MacDorman K. F.: Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the Uncanny Valley. In *Proc. of ICCS/CogSci Long Symp.: Towards Social Mechanisms of Android Science* (2006). 7

[Mat03] Mateas M.: Expressive ai: Games and artificial intelligence. In *DiGRA Conference* (2003). 15

[MBB12] McDonnell R., Breidt M., Bülthoff H. H.: Render me real? Investigating the effect of render style on the perception of animated virtual humans. *ACM Transaction on Graphics 31*, 4 (2012), 91:1–91:11. 7, 8

[MBF17] Mader B., Banks M. S., Farid H.: Identifying computer-generated portraits: The importance of training and incentives. *Perception 46*, 9 (2017), 1062–1076. doi:10.1177/0301006617713633. 8

[MBZ*16] Matsuyama Y., Bhardwaj A., Zhao R., Romeo O., Akoju S., Cassell J.: Socially-Aware Animated Intelligent Personal Assistant Agent, dec 2016. doi:10.18653/v1/w16-3628. 20

[McC93] McCloud S.: *Understanding Comics*. William Morrow Paperbacks, 1993. 7

[MDCO06] McDonnell R., Dobbyn S., Collins S., O'Sullivan C.: Perceptual evaluation of lod clothing for virtual humans. In *Proc. of Symp. on Computer Animation* (2006), SCA '06, pp. 117–126. 9

[Meh80] Mehrabian A.: *Basic Dimensions for a General Psychological Theory Implications for Personality, Social, Environmental, and Developmental Studies*. University of California, 1980. 12

[MGCPV07] Morrison E. R., Gralewski L., Campbell N., Penton-Voak I. S.: Facial movement varies by sex and is related to attractiveness. *Evolution and Human Behavior 28*, 3 (2007), 186–192. 12

[MGHK09] MacDorman K. F., Green R. D., Ho C.-C., Koch C. T.: Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior 25*, 3 (2009), 695–710. 9, 10

[MGM05] Maatman R. M., Gratch J., Marsella S.: Natural Behavior of a Listening Agent. In *Intelligent Virtual Agents, 5th International Working Conference* (Kos, Greece, sep 2005), Panayiotopoulos T., Gratch J., Aylett R., Ballin D., Olivier P., Rist T., (Eds.). 19

[MGP06] Mehl M. R., Gosling S. D., Pennebaker J. W.: Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of personality and social psychology 90*, 5 (2006), 862. 14

[MJH*09] McDonnell R., Jörg S., Hodgins J. K., Newell F., O'sullivan C.: Evaluating the effect of motion and body shape on the perceived sex of virtual characters. *ACM Transactions on Applied Perception (TAP) 5*, 4 (2009), 20. 12

[MJM*09] McDonnell R., Jörg S., McHugh J., Newell F. N., O'Sullivan C.: Investigating the role of body shape on the perception of emotion. *ACM Transactions on Applied Perception (TAP) 6*, 3 (2009), 14. 13

[MLD*08] McDonnell R., Larkin M., Dobbyn S., Collins S., O'Sullivan C.: Clone attack! perception of crowd variety. *ACM Transaction on Graphics (TOG) 27*, 3 (2008), 26. 14

[MLH*09] McDonnell R., Larkin M., Hernandez B., Rudomin I., , O'Sullivan C.: Eye-catching Crowds: saliency based selective variation. *ACM Transaction on Graphics 28*, 3 (2009), 55:1 – 55:10. 11

[MMG11] McNamara A., Mania K., Gutierrez D.: Perception in graphics, visualization, virtual environments and animation. In *SIGGRAPH Asia 2011 Courses* (2011), pp. 17:1–17:137. 2

[MO10] McDonnell R., O'Sullivan C.: Movements and voices affect perceived sex of virtual conversers. In *Proc. of Symp. on Applied Perception in Graphics and Visualization* (2010), ACM, pp. 125–128. 12

[MR08] Mania K., Reinhard E.: Perceptually-motivated graphics. In *Eurographics Tutorials* (2008). 2

[MXH*10] Ma W., Xia S., Hodgins J. K., Yang X., Li C., Wang Z.: Modeling style and variation in human motion. In *Proc. of Symp. on Computer Animation* (2010), Eurographics Association, pp. 21–30. 14

[NHMIK00] Niedenthal P. M., Halberstadt J. B., Margolin J., Innes-Ker Å. H.: Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology 30*, 2 (2000), 211–222. 5

[Now01] Nowak K.: The influence of anthropomorphism on social judgment in social virtual environments. In *Annual Convention of the International Communication Association, Washington, DC* (2001). 15

[NR11] Nitzsche B., Rose K.: *Bodyshop: The Photoshop Retouching Guide for the Face and Body*. Wiley, 2011. 10

[NWAW10] Neff M., Wang Y., Abbott R., Walker M.: Evaluating the effect of gesture and language on personality perception in conversational agents. In *Int. Conf. on Intelligent Virtual Agents* (2010), Springer, pp. 222–235. 14

[OB10] Olkkonen M., Brainard D. H.: Perceived glossiness and lightness under real-world illumination. *Journal of Vision 10*, 9 (2010). 10

[OB11] Olkkonen M., Brainard D. H.: Joint effects of illumination geometry and object shape in the perception of surface reflectance. *i-Perception 2* (2011), 1014–1034. 10

[OHM*04] O'Sullivan C., Howlett S., Morvan Y., McDonnell R., O'Conor K.: Perceptually adaptive graphics. In *Eurographics STAR* (2004). 2

[PCWA*18] Pan X., Collingwoode-Williams T., Antley A., Brenton H., Congdon B., Drewett O., Gillies M. F., Swapp D., Pleasence P., Fertleman C., et al.: A study of professional awareness using immersive virtual reality: the responses of general practitioners to child safeguarding concerns. *Frontiers in Robotics and AI 5* (2018), 80. 20

[PGB*12a] Pan X., Gillies M., Barker C., Clark D. M., Slater M.: Socially anxious and confident men interact with a forward virtual woman: an experimental study. *PloS one 7*, 4 (jan 2012), e32931. doi:10.1371/journal.pone.0032931. 17, 18, 19, 20

[PGB*12b] Pan X., Gillies M., Barker C., Clark D. M., Slater M.: Socially anxious and confident men interact with a forward virtual woman: an experimental study. *PloS one 7*, 4 (2012). 19

[PGS15] Pan X., Gillies M., Slater M.: Virtual Character Personality Influences Participant Attitudes and Behavior - An Interview with a Virtual Human Character about Her Social Anxiety, feb 2015. 18, 20

[PH15] Pan X., Hamilton A. F. d. C.: Automatic imitation in a rich social context with virtual characters. *Frontiers in psychology 6* (2015), 790. 19

[PHKD00] Plant E. A., Hyde J. S., Keltner D., Devine P. G.: The gender stereotyping of emotions. *Psychology of Women Quarterly 24*, 1 (2000), 81–92. 13

[Plu01] Plutchik R.: The nature of emotions human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist 89*, 4 (2001), 344–350. 12

[PO11] Pražák M., O'Sullivan C.: Perceiving human motion variety. In *Proc. of Symp. on Applied Perception in Graphics and Visualization* (2011), ACM, pp. 87–92. 14

[Pol15] Poland J. L.: *Light, Camera, Emotion! An Examination on Film Lighting and It's Impact on Audiences' Emotional Response*. Cleveland State University, 2015. 11

[PPBS01] Pollick F. E., Paterson H. M., Bruderlin A., Sanford A. J.: Perceiving affect from arm movement. *Cognition 82*, 2 (2001), B51–B61. 12

[PS16] Pan Y., Steed A.: A comparison of avatar-, video-, and robot-mediated interaction on users trust in expertise. *Frontiers in Robotics and AI 3* (2016), 12. doi:10.3389/frobt.2016.00012. 8

[PSB*16] Pan X., Slater M., Beacco A., Navarro X., Rivas A. I. B., Swapp D., Hale J., Forbes P. A. G., Denvir C., Hamilton A. F. d. C., et al.: The responses of medical general practitioners to unreasonable patient demand for antibiotics-a study of medical ethics using immersive virtual reality. *PloS one 11*, 2 (2016). 20

[PtP06] Pont S. C., te Pas S. F.: Material – illumination ambiguities and the perception of solid objects. *Perception 35* (2006), 1331–1350. 10

[Rad02] Rademacher P. M.: *Measuring the Perceived Visual Realism of Images*. PhD Thesis, 2002. 11

[RCB05] Ritchie K., Callery J., Biri K.: *The Art of Rigging: A Definitive Guide to Character Technical Direction with Alias Maya*, vol. 1. Alias, 2005. 7

[RFJ18] Raiturkar P., Farid H., Jain E.: *Identifying Computer-Generated Portraits: an Eye Tracking Study*. Tech. rep., University of Florida, 08 2018. 11

[RGKK*14] Rivera-Gutierrez D. J., Kopper R., Kleinsmith A., Cendan J., Finney G., Lok B.: Exploring gender biases with virtual patients for high stakes interpersonal skills training. In *International Conference on Intelligent Virtual Agents* (2014), Springer, pp. 385–396. 20

[Rho06] Rhodes G.: The evolutionary psychology of facial beauty. *Annu. Rev. Psychol. 57* (2006), 199–226. 14

[RJD*07] Raij A. B., Johnsen K., Dickerson R. F., Lok B. C., Cohen M. S., Duerson M., Pauly R. R., Stevens A. O., Wagner P., Lind D. S.: Comparing interpersonal interactions with a virtual human to those with a real human. *IEEE transactions on visualization and computer graphics 13*, 3 (2007), 443–457. 20

[RLCW01] Rademacher P. M., Lengyel J., Cutrell E., Whitted T.: Measuring the perception of visual realism in images. In *Proc. of the Eurographics Symp. on Rendering (EGSR)* (2001), pp. 235–247. 11

[RPA*15] Ruhland K., Peters C. E., Andrist S., Badler J. B., Badler N. I., Gleicher M., Mutlu B., McDonnell R.: A review of eye gaze in virtual agents, social robotics and hci: Behaviour generation, user interaction and perception. *Computer graphics forum 34*, 6 (2015), 299–326. 12

[RSSF12] Rauthmann J. F., Seubert C. T., Sachse P., Furtner M. R.: Eyes as windows to the soul: gazing behavior is related to personality. *Journal of Research in Personality 46*, 2 (2012), 147–156. 14

[Rus80] Russell J. A.: A circumplex model of affect. *Journal of personality and social psychology 39*, 6 (1980), 1161. 12

[RZM15] Ruhland K., Zibrek K., McDonnell R.: Perception of personality through eye gaze of realistic and cartoon models. In *Proc. of Symp. on Applied Perception* (2015), ACM, pp. 19–23. 14

[SB12] Segovia K. Y., Bailenson J. N.: Virtual imposters: Responses to avatars that do not look like their controllers. *Social Influence 7*, 4 (2012), 285–303. doi:10.1080/15534510.2012.670906. 9

[SBC*12] Schröder M., Bevacqua E., Cowie R., Eyben F., Gunes H., Heylen D., Ter Maat M., McKeown G., Pammi S., Pantic M., Pelachaud C., Schuller B., De Sevin E., Valstar M., Wöllmer M.: Building autonomous sensitive artificial listeners. *IEEE Transactions on Affective Computing 3*, 2 (2012), 165–183. doi:10.1109/T-AFFC.2011.34. 19

[Sch11] Schröder M.: The semaine api: A component integration framework for a naturally interacting and emotionally competent embodied conversational agent, 2011. 15

[SJ16] Schwind V., Jäger S.: The uncanny valley and the importance of eye contact. *i-com* (2016), 93–104. 11

[Sla09] Slater M.: Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos Trans R Soc Lond B Biol Sci 364*, 1535 (dec 2009), 3549–3557. doi:10.1098/rstb.2009.0138. 16

[SN07] Seyama J., Nagayama R. S.: The Uncanny Valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments 16*, 4 (2007), 337–351. 9

[SPBC06] Slater M., Pertaub D.-P., Barker C., Clark D. M.: An Experimental Study on Fear of Public Speaking Using a Virtual Environment. *CyberPsychology & Behavior* (2006). doi:10.1089/cpb.2006.9.627. 17, 18

[SPS09] Schyns P. G., Petro L. S., Smith M. L.: Transmission of facial expressions of emotion co-evolved with their efficient decoding in the brain: behavioral and brain evidence. *PLoS One 4*, 5 (2009), e5625. 13

[SS02] Slater M., Steed A.: Meeting people virtually: Experiments in shared virtual environments. In *The Social Life of Avatars*. Springer, 2002, pp. 146–171. 15

[SSR11] Schwarzkopf D. S., Song C., Rees G.: The surface area of human v1 predicts the subjective experience of object size. *Nature neuroscience 14*, 1 (2011), 28. 4

[SWY07] Schneider E., Wang Y., Yang S.: Exploring the Uncanny Valley with Japanese video game characters. In *Situated Play: Proc. of Digital Games Research Association Conference* (2007). 7

[TNC13] Tinwell A., Nabi D. A., Charlton J. P.: Perception of psychopathy and the uncanny valley in virtual characters. *Computers in Human Behavior 29*, 4 (2013), 1617–1625. 14

[TOdHCD06] Thalmann D., O'Sullivan C., de Heras Ciechomski P., Dobbyn S.: Populating virtual environments with crowds. In *Eurographics Tutorials* (2006). 9

[TOY*07] Thalmann D., O'Sullivan C., Yersin B., Maim J., McDonnell R.: Populating virtual environments with crowds. In *Eurographics Tutorials* (2007). 2, 9

[TPS*18] Thaler A., Piryankova I., Stefanucci J. K., Pujades S., de la Rosa S., Streuber S., Romero J., Black M. J., Mohler B. J.: Visual perception and evaluation of photo-realistic self-avatars from 3D body scans in males and females. *Frontiers in ICT 5* (Sept. 2018), 1–14. 9

[TSJR12] Talbot T. B., Sagae K., John B., Rizzo A. A.: Sorting out the virtual patient: how to exploit artificial intelligence, game technology and sound educational practices to create engaging role-playing simulations. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS) 4*, 3 (2012), 1–19. 20

[Van09] Vangorp P.: *Human Visual Perception of Materials in Realistic Computer Graphics*. PhD Thesis, University Leuven, 2009. 10

[VBC*16] Volante M., Babu S. V., Chaturvedi H., Newsome N., Ebrahimi E., Roy T., Daily S. B., Fasolino T.: Effects of virtual human appearance fidelity on emotion contagion in affective interpersonal simulations. *IEEE Transaction on Visualization and Computer Graphics 22*, 4 (2016), 1326–1335. 13

[VGS*06] Vinayagamoorthy V., Gillies M., Steed A., Tanguy E., Pan X., Loscos C., Slater M.: Building Expression into Virtual Characters. In *Eurographics Conference State of the Art Reports* (2006). 2

[VHPM13] Verberne F. M., Ham J., Ponnada A., Midden C. J.: Trusting digital chameleons: The effect of mimicry by a virtual social agent on user trust. In *International Conference on Persuasive Technology* (2013), Springer, pp. 234–245. 19

[VLD07] Vangorp P., Laurijssen J., Dutre P.: The influence of shape on the perception of material reflectance. *ACM Transaction on Graphics 26*, 3 (2007), 77:1–77:9. 10

[VMD*14] Volkova E. P., Mohler B. J., Dodds T. J., Tesch J., Bülthoff H. H.: Emotion categorization of body expressions in narrative scenarios. *Frontiers in Psychology 5* (2014). 12

[WBC*07] Wallraven C., Bülthoff H. H., Cunningham D. W., Fischer J., Bartz D.: Evaluation of real-world and computer-generated stylized facial expressions. *ACM Transactions on Applied Perception 4*, 3 (2007). 7, 8, 13

[WBCB08] Wallraven C., Breidt M., Cunningham D. W., Bülthoff H. H.: Evaluating the perceptual realism of animated facial expressions. *ACM Transactions on Applied Perception 4*, 4 (2008), 4:1–4:20. 7

[WBR16] Washburn M., Bordnick P., Rizzo A. Â.: A pilot feasibility study of virtual patient simulation to enhance social work studentsâĂŹ brief mental health assessment skills. *Social work in health care 55*, 9 (2016), 675–693. 20

[WDM16] Wisessing P., Dingliana J., McDonnell R.: Perception of lighting and shading for animated virtual characters. In *Proc. of ACM Symp. of Applied Perception (SAP)* (2016), pp. 25–29. 11

[Wer23] Wertheimer M.: A brief introduction to gestalt, identifying key theories and principles. *Psychol Forsch 4* (1923), 301–350. 4

[WG10] Wang N., Gratch J.: Don't just stare at me! In *Conference on Human Factors in Computing Systems - Proceedings* (2010), vol. 2, pp. 1241–1249. doi:10.1145/1753326.1753513. 19

[WGH13] Wang Y., Geigel J., Herbert A.: Reading personality: Avatar vs. human faces. In *Humaine Association Conference on Affective Computing and Intelligent Interaction* (2013), pp. 479–484. doi:10.1109/ACII.2013.85. 9

[WGR*18] Waltemate T., Gall D., Roth D., Botsch M., Latoschik M. E.: The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Transactions on Visualization and Computer Graphics 24*, 4 (2018), 1643–1652. doi:10.1109/TVCG.2018.2794629. 8, 9

[YBR07] Yee N., Bailenson J. N., Rickertsen K.: A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems* (2007), CHI '07, pp. 1–10. doi:10.1145/1240624.1240626. 8

[ZAJ*15] Zell E., Aliaga C., Jarabo A., Zibrek K., Gutierrez D., McDonnell R., Botsch M.: To stylize or not to stylize?: The effect of shape and material stylization on the perception of computer-generated faces. *ACM Transactions on Graphics 34*, 6 (2015), 184:1–184:12. 7, 8, 10, 11

[ZHRM15] Zibrek K., Hoyet L., Ruhland K., McDonnell R.: Exploring the effect of motion type and emotions on the perception of gender in virtual humans. *ACM Transactions on Applied Perception (TAP) 12*, 3 (2015), 11. 12, 13, 14

[ZKM17] Zibrek K., Kokkinara E., McDonnell R.: Don't stand so close to me: investigating the effect of control on the appeal of virtual humans using immersion and a proximity-based behavioral task. In *Proceedings of the ACM Symposium on Applied Perception* (2017), pp. 1–11. 15

[ZKM18] Zibrek K., Kokkinara E., McDonnell R.: The effect of realistic appearance of virtual characters in immersive environments-does the character's personality play a role? *IEEE Transactions on Visualization and Computer Graphics 24*, 4 (2018), 1681–1690. 4, 13, 15

[ZM14] Zibrek K., McDonnell R.: Does render style affect perception of personality in virtual humans? In *Proc. of ACM Symp. on Applied Perception (SAP)* (2014), pp. 111–115. 14

[ZSBC16] Zhao R., Sinha T., Black A. W., Cassell J.: Socially-aware virtual agents: Automatically assessing dyadic rapport from temporal patterns of behavior. In *International conference on intelligent virtual agents* (2016), Springer, pp. 218–233. 15

[ZVC96] Zebrowitz L. A., Voinescu L., Collins M. A.: "wideeyed" and "crooked-faced": Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin 12*, 12 (1996), 1258–âĂŞ1269. 9