The role of eye gaze in virtual characters on their perceived believability, social presence and the application of Theory of Mind

by

Z. Garama

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Monday November 18, 2019 at 08:30 AM.

Student number:	4217675	
Project duration:	September 10, 2018 – November 18, 2019	
Thesis committee:	Dr. Ir. Willem-Paul Brinkman,	TU Delft, supervisor
	Ding Ding,	TU Delft, daily supervisor
	Prof. Dr. Mark A. Neerincx,	TU Delft
	Dr. Klaus Hildebrandt,	TU Delft

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Abstract

Making virtual characters seem sentient is the main goal of the study of believable virtual characters. Unbelievable behaviour, especially eye gaze behaviour, could make the whole virtual character seem less believable. In human-human social interactions eye gaze is used to interact non-verbally. People also apply Theory of Mind (ToM) to reason about other people's mental states. If a virtual character were to respond to a person's gazing behaviour with its own gazing behaviour, will it seem more believable and socially present to this person? Do people look at- and follow the gaze of such virtual characters more? Do people apply ToM to reason about the mental states of such virtual characters? This study employed a within-subject design (n=24) with two conditions that were given at the same moment. The conditions were represented by two virtual characters where one had gaze aware behaviour and the other did not. Using a think-aloud procedure during and questionnaires after the Virtual Reality experience the believability, social presence and the application of ToM were assessed. Gaze behaviour with respect to the virtual characters was assessed as well using the onboard eye-tracker of the Virtual Reality Head Mounted Display. The results suggest that virtual characters were more believable and socially present when they respond to people's gaze behaviour through their own gaze behaviour compared to a virtual character without such gaze behaviour. Results also showed that in some situations people looked more at the virtual character without gaze aware behaviour than at the virtual character with such behaviour. However, gaze following behaviour and whether or not people apply ToM were inconclusive. In conclusion, virtual characters that respond to people's gaze behaviour through their own gaze behaviour are perceived as more believable and more socially present than virtual characters without such gaze aware behaviour.

Preface

This document is the result of my graduation thesis to obtain my master's degree in Computer Science. I would not be here without the help of a number of people.

First of all, I want to thank my participants for their critical role in the experiment. Many thanks as well to Mark Neerincx and Klaus Hildebrandt for being part of this committee. Making sure I can graduate and giving me tips on the defense. I also want to thank Tore Knabe for providing an awesome Unity plugin which helped me immensely in creating the right gazing behaviour for my virtual characters. Thank you, Ding and Willem-Paul for guiding me through the process. Your advice made sure I delivered the best quality work. Willem-Paul, your enthiousiasm and openness makes communication a joy. With your experience and critical thinking you challenged me to get out of my comfort zone because you knew I could do it. Ding, your hard working and sense of responsibility inspires me. You have pushed me to provide quality work that we can both be proud of and I know you will continue to do that while we try to get this study published together. I also want to thank the other master students, Pascal, Siyu (Sylvia), Mo, David, Mitchell, Han, Salim and Jeffrey, who have been supervised by Willem-Paul. You stood with me during the weekly meetings, offering invaluable insights. I am gratefull for Lars (van der Zwan), Marnix and Jamey for keeping me company in the INSYGHT lab while we were all working on our thesis. There have been many funny moments where we were joking around, which I will always look back to with nostalgia (remember Henk and Yari forever). Wessel, you quicly became one of my best friends. You are working hard on self-improvement and you have inspired me to work harder on improving myself as well. Last but definitely not least, I would like to thank my family and closest friends, Annelies and Vera, for not only supporting me during the thesis but also during the phases of my life that lead up to it. I am particularly gratefull for my boyfriend, Lars. Your love and belief in me was indispensable. You cheered for me and supported me every step of the way. You are amazing.

Z. Garama
Delft, November 2019

Contents

1 Introduction

Making the unaware aware, can you look a virtual character in the eye and believe it is sentient? Virtual characters can look very realistic but their gazing behaviour often seems simple. As eye gaze is a very important medium for social signals in human-human interaction, an unrealistic implementation of a virtual characters' gaze behaviour has a negative impact on the overall perception of the virtual character [23, 76]. This study explores the impact of eye gaze on whether people perceive a virtual character as a person.

Without gazing behaviour virtual characters might lack the ability to communicate their mental states properly. Eye gaze behaviour is a cognitively special social signal with unique, hard-wired pathways in the brain dedicated to their interpretation [28]. People can not read other peoples' minds but they can make guesses based on that persons' behaviour. This ability to estimate other people's mental states, i.e. their beliefs, desires and intentions, based on non-verbal cues is called Theory of Mind (ToM) [28]. Eye gaze plays a critical role in the development of this ability [8, 70] and eye gaze remains critical even later in life when someone else's intentions are ambiguous [64]. Furthermore it is hard to read complex emotions and mental states without looking at the eyes [9, 12, 26, 27, 37, 59, 60].

Machine Consciousness is a field of research that aims at understanding, simulating and replicating human consciousness in machines [31]. It models the behaviour of virtual characters after psychological theories of human cognition with the goal of making them more believable [4]. Bates et. al. [11], from the perspective of a combination of autonomous agents and believable characters from the traditional character-based arts such as animation, film or literature, defines the main goal of believable virtual characters to give an illusion of life. It is not their physical appearance that is important for this illusion but their behaviour has to seem natural, appealing, life-like [50]. In actuality this way of thinking seems the wrong way around. Humans will interpret almost anything with self-propelled motion, or anything that makes a non-random sound, as an agent with goals and desires [8, 16]. This will, of course, lead to some false positives but according to Baron-Cohen [8] experience could then let that person override the signal that there is an agent, given other types of knowledge. In that sense, the goal is not to create life-like virtual characters per se but to remove or improve the elements of a virtual character that make it unbelievable.

Social presence is a reflection of the degree to which one believes that he or she is in the presence of, and interacting with, other human beings [39]. Thus, for a virtual character to be perceived as a person, people need to perceive the virtual character as socially present. Eye gaze seems to be important for this feeling as a lack of eye contact has been shown to reduce social presence [6, 74].

To sum up, our vision is that virtual characters could seem more like people if they were able to communicate socially using eye gaze. We think that these virtual characters would become more believable and more socially present as well as have more mental states attributed to them than virtual characters without this behaviour.

2

Related work

2.1. Eye gaze as a social signal

Eye gaze can convey different social signals. The most commonly reported social signals are: focus of attention [3, 44], sign of intent [17], joint attention [79], desire/avoidance for interaction [79], regulating the flow of conversation [56], and establishment and recognition of social relationships [44, 79]. From these, some of the more fundamental social signals are *focus of attention* and *joint attention*. This is because they depend on eye gaze only and not on more complex interactions with other social behaviour such as speech in the case of regulating the flow of conversation or social context in the case of establishment and recognition of social relationships. These signals are so fundamental that they are learned very early in a human child's development [8]. By the age of three months, infants display maintenance (and thus recognition) of eye contact, by the age of nine months children begin to exhibit gaze following, and by the age of eighteen months children will follow gaze outside their field of view [8].

Eye gaze functions as a direct proxy for a person's attention [7]. Someone's focus of attention can be perceived by others by combining that person's gaze direction with knowledge of probable targets of interest in the world [48, 75]. Focus of attention is a crucial learning element in humans [36] and virtual characters that make use of this knowledge have been shown to be able to make a learner pay more attention, be more engaged in learning activities and guide a learners' attention through material in a thoughtful way [62, 68, 80].

Mutual gaze, also called eye contact, is a subset of the social signal *focus of attention* and it is different from the other social signals in that it is inherently reciprocal. It not only indicates that your attention is focused on the other person, but also that that person's attention is focused on you. That is probably why being gazed at is such a powerful attention grabbing stimulus, even if it is done by virtual characters [61]. However, offsets between screens and cameras can make it difficult or even impossible to establish mutual gaze [74]. Using Virtual Reality (VR) with inbuilt eye-trackers mitigates this problem as the mapping is much more direct. This is important because presence plays a major role in VR and a lack of mutual gaze can reduce social presence [6, 74].

In humans, joint attention is critical for the development of Theory of Mind [8] as well as other things [70]. It is the sum of two sequential attentional processes: following the line of sight of another person and then focusing on the same attended object [28]. It serves as the initial mechanism for infants to share experiences with others and to negotiate shared meanings of objects and events in the world [70]. The link between joint attention and ToM has been studied by researchers in a variety of fields. Philosophers have been interested in joint attention as a precursor to a theory of other minds [22, 82]. Evolutionary psychologists and primatologists have focused on the evolution of these simple social skills throughout the animal kingdom as a means of evaluating both the presence of ToM and as a measure of social functioning [38, 66, 67]. From the perspective of creating human-like robots that exhibit social skills, joint attention has been studied by Artificial Intelligence (AI) researchers as a way to communicate naturally with humans [70]. They also argued that it would allow the robot to express its mental states through social interactions without relying upon vocabulary [70].

2.2. Believability

What happens when virtual characters are unbelievable? To answer this question one can look at the Uncanny Valley hypothesis that proposes that by increasing the human-like appearance of a robot, the affinity for it increases as well [55, 84]. However, when a robot's appearance becomes sufficiently human-like but is still distinguishable from real humans, people's emotional response becomes very negative. Then, once the appearance of a robot becomes indistinguishable from a real human the affinity reaches its optimum at the same level as for humans. According to the original paper by Marashiro Mori [55], dead humans such as corpses and zombies are the most eery and were placed deep into the Uncanny Valley. This relation between death and the Uncanny Valley can also be found in reports from the cinematic game Heavy Rain [25]. The designers failed to create believable behaviour even though the visual realism was high. Players reported on the inanimacy of the eyes, calling them awkward, bizarre and glazed-over [23]. Similarly, eyes from virtual characters in the game Alias have been described as "monsters with dead eyes" [76].

For agents to become truly believable, the eyes are especially important to get right. For example, Brenton et. al. [15] argued that the Uncanny Valley is probably related to our innate ability to extract social and emotional information from human faces. People are highly attuned to the cues given by faces, and therefore find it easier to spot small variations [24]. The eyes are especially powerful in our recognition of life because they allow us to attribute mental states to another organism [81]. Perceptual cues that indicate falsehood are thus especially potent in the eyes [15].

Virtual characters using eye gaze to communicate social signals might be able to decrease issues with the Uncanny Valley. This is because where the eye gaze fails to communicate intent, the ensuing unpredictable behaviour promotes fear [15]. Adding to that is the issue that higher graphical realism suggests that the virtual character is a person which raises high expectations for motion and behaviour. When these expectations are not met, this suggests that the virtual character is not real, creating a perceptual paradox which has been suggested as a potential cause for the Uncanny Valley [15, 73]. These issues suggest that a promising way to improve the believability of virtual characters is to have virtual characters express social behaviour using eye gaze. This could help with communication, decrease unpredictable behaviour and match expectations better.

2.3. Theory of Mind (ToM)

Eye gaze is essential in social situations because it helps other people in several different ways to apply ToM. First and foremost, it helps in disambiguating and understanding intentions. In fact, it is the first place people look for to disambiguate conflicting actions [64]. This is especially important in social situations as actions can often be interpreted in many different ways. This is one of the main reasons why people with Autism Spectrum Disorder (ASD) have such a hard time in social situations [8]. Second, although people can read simple emotions from the lower half of the face, the eyes are necessary to recognise complex emotions. A study by Baron-Cohen [9] showed that for basic emotions, the whole face is more informative than either the eyes or the mouth alone. However, for the complex mental states, seeing the eyes alone produced significantly better performance than seeing the mouth alone, and was as informative as seeing the whole face. Other studies have also shown the role of the eyes in the recognition of complex emotions (scheming, admiration, interest, thoughtfulness, deceit, fear, surprise, etc.) [12, 26, 27, 37, 59, 60]. Third, it plays a key role in personality evaluation and person perception [12, 40, 45, 53, 54]. Eye gaze information helps in evaluating personalities as categorisation is quicker in direct gaze conditions compared to deviated or closed conditions [51]. It also makes it easier to recognise and remember faces [40].

2.4. Social presence

Social presence is defined by Heeter [39] as a reflection of the degree to which one believes that he or she is in the presence of, and interacting with, other human beings. The implicit, psychological question is “What properties of humans elicit attributions of cognitive states to representations, as if those representations contained minds?” [13]. Biocca [13] argued that all the dimensions of social presence used by researchers circle one basic phenomenon: that social presence may be the product of the process of applying ToM to a representation. When interacting with virtual characters users apply ToM and respond socially, even when they know that no mind or social other really exists [13]. Fundamentally, when responding to all social representations, people know that the other is just patterns of light on a screen, yet the social responses are automatic [8, 16, 57, 58]. Biocca [13] also proposed that social presence theory may benefit by seeking to forge a deeper link between the brain, the properties that apply ToM in representations; and technology, the properties that simulate agency in inanimate things such as pixels, paint, and clay. Could social presence be the relation between applying ToM to representations and eye gaze as a property that simulates believability in animate things such as virtual characters?

2.5. Hypotheses

Given previous research, we predict that creating virtual characters with socially aware eye gaze has a positive effect on the perception of virtual characters. These virtual characters will be called gaze aware virtual characters (GAVC) from now on. This prediction is formulated in the following hypotheses:

H1: People look at - and follow the gaze of - a gaze aware virtual character more than they do with a non- gaze aware virtual character.

H2: People perceive a gaze aware virtual character as more believable and more socially present than a non- gaze aware virtual character.

H3: People employ Theory of Mind more often to reason about the mental state of a gaze aware virtual character than they do to reason about the mental state of a non- gaze aware virtual character.

3
System

To examine the hypotheses, a virtual environment and two virtual characters with their respective gazing behaviour were developed. It was an immersive 3D Virtual Reality system mainly designed to provide a waiting room scenario.

3.1. Design of the virtual environment

The system assumed that people were seated in real life, so in the virtual environment the camera was anchored to a chair. This made it seem like people were seated in the virtual environment as well even though they had no bodily representation.

Waiting room

A waiting room scenario was chosen because people usually keep an eye on each other but do not talk, they generally have no relation to each other, often contain many objects, and it can be quite a small room. Figure 3.1 shows the virtual waiting room. The person was seated at a table roughly in the middle of the room. At the other side of the table two virtual characters were seated. The virtual characters were placed at a maximum distance from the person such that their eyes and their movements were still clearly visible. On the table several items were placed which allowed people to see both the virtual characters and objects at the same time. The other objects were scattered around the room and needed more head movement to see. This gave different opportunities for people to study the behaviour of the virtual characters. Two objects were placed behind the person so that they could be discovered later on by following the gaze of the virtual characters. No objects were placed behind the virtual characters because that would require the virtual characters to turn around in their chair, which is more body movement than eye or head movement. All objects in the room were chosen with the purpose in mind that the person could create stories about the objects. This allowed people to keep their minds busy and prevented them from getting distracted by random thoughts. More importantly, it ensured that they kept looking around as joint attention per definition requires an object to share attention over.

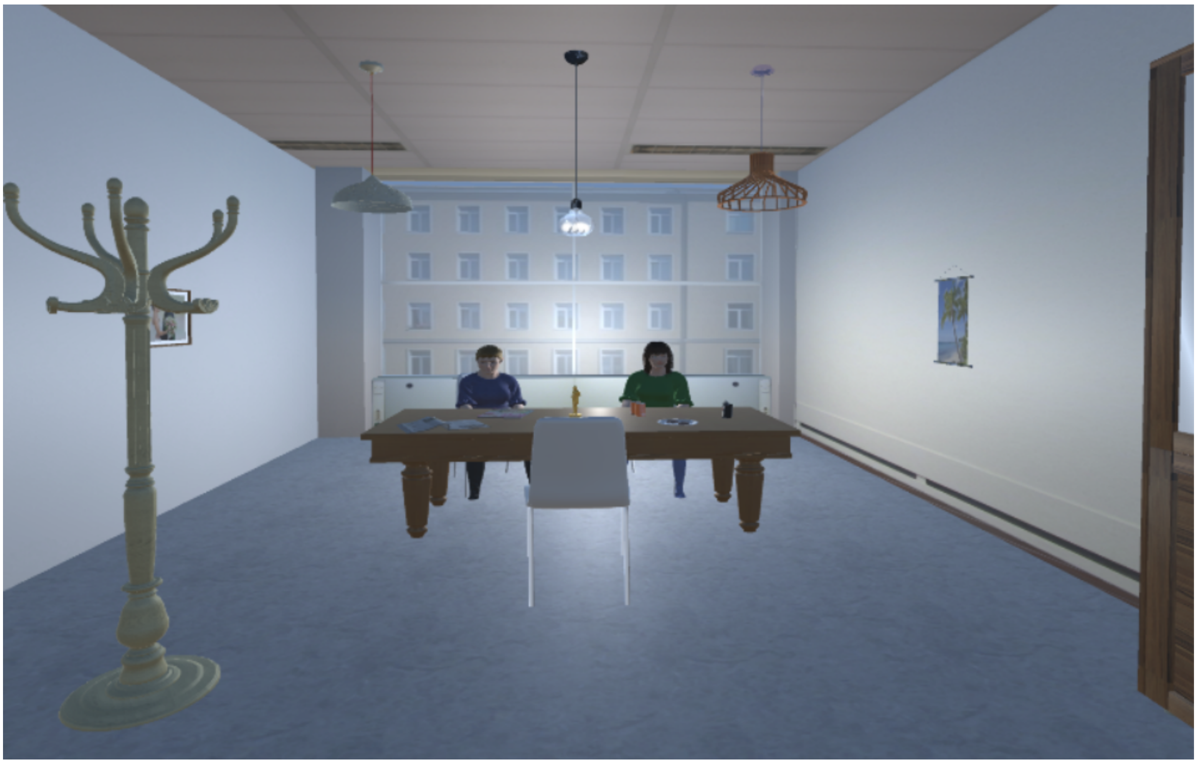


Figure 3.1: The virtual waiting room

3.2. Design of the virtual characters

Two sets of virtual characters were designed, one set was female and the other was male (see figure 3.2 and figure 3.3 respectively). The virtual characters were gender matched.

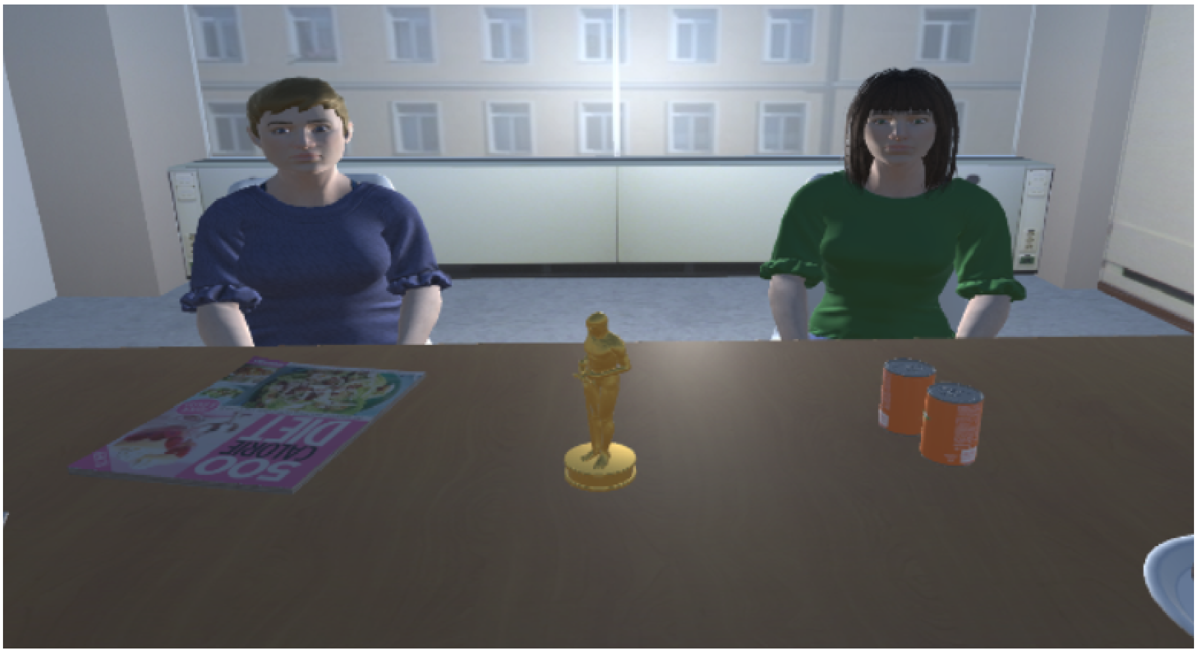


Figure 3.2: Virtual characters when participant is female

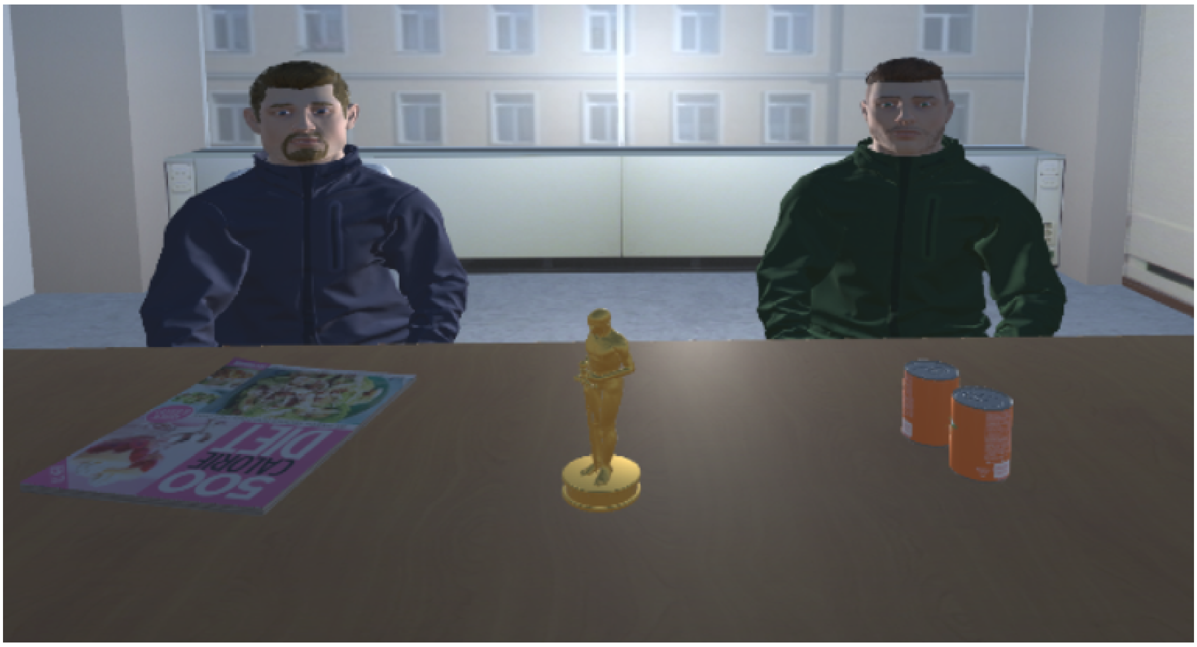


Figure 3.3: Virtual characters when participant is male

Gaze behaviour

In order to create realistic looking gazing behaviour for both virtual characters Tore Knabe's "Realistic Eye Movements" [47] Unity asset was used. This asset used models that were based on video's of real people's eye and head movements [42, 46, 49] to make the eye and head movements of the virtual characters as realistic as possible.

When a person looked at an object in the room, the gaze aware virtual character (GAVC) followed their gaze to that object (joint attention). In code this worked by "pulling" the gaze of this virtual character to the object that a person is looking at at that moment. A small pilot study was conducted with 4 participants (4 male, age 24-27) to determine the threshold for when an object or virtual character could be considered looked at. This was to prevent a glance over an object to be confused with looking at the object. People were asked to think aloud during the study. The initial threshold was set to 1 second. However, in some cases the system did not detect a person looking at an object when they had clearly mentioned the object. Therefore the threshold was lowered to 0.7 seconds. Another thing that was important to take into account was when the person looked at different objects relatively quickly. In that case the GAVC would finish looking at the first one and would then move on to the object that the person was currently looking at. Also, when a person looked directly at the GAVC, the virtual character would look back and establish eye contact. The duration of the eye contact depended on the person's gender and was based on a study by Mark Cook [20], 0.9 seconds for males and 1.4 seconds for females.

The non- gaze aware virtual character acted independently from other people's gaze behaviour. It looked around the room in an idle fashion. People's gaze tends to focus on objects instead of at random points in space [63]. That is why this virtual character looked sequentially at objects chosen randomly from a list of objects in the room.

4

Method

4.1. Experimental design

The experiment had a within-subject design with two conditions that were given at the same moment. The conditions were represented by two virtual characters where one had gaze aware behaviour and the other did not. Which condition the virtual character was assigned to (gaze aware or control) was determined dynamically at the start of the experiment and was randomised across the set of people. The dependent variables aimed to measure the believability, social presence, gaze behaviour and the application of Theory of Mind. The experiment was divided into two phases. During both phases the participant was seated across from the two virtual characters in a waiting room. The participants could not move around the room and could not noticeably interact with the objects in the room. It was only possible for them to look around.

In phase 1, the GAVC followed the gaze of the participant to objects in the room (joint attention). When a participant looked at the GAVC directly the virtual character would establish eye contact (mutual gaze). The control virtual character (CVC) looked around idly. The duration of this phase was 2.5 minutes, after which phase 2 automatically started.

Phase 2 was the measuring phase and had a duration of 1.5 minutes. Here, both virtual characters had the same behaviour. This means that the GAVC stopped following the gaze of the participant to objects. Instead it looked idly at the objects in the room. Virtual characters did not establish eye contact in this phase. The expectation was that the participant was more interested in the behaviour of the GAVC because of phase 1. That is why in phase 2 the expectation was that the gaze of the virtual character to objects in the scene would be perceived by the participant as an initiation of joint attention. As a result the participant could react to the joint attention initiation and start following the gaze of the GAVC. The shift between the phases could only be noticed by a participant by noticing the changes in the virtual characters' behaviour. The phase shift had not been made explicit because the participants should not have changed their behaviour consciously.

The duration of these phases was determined using a small pilot study with 4 participants (4 male, age 22-25). In this pilot study participants were asked to think aloud as well. An observer noted how long it took for a participant to mention noticing a difference between the virtual characters. How long it took participants to really figure out that the gazing behaviour of one of the virtual characters was dependent on their own gazing behaviour was also noted down. The medians of these two metrics per participant was then averaged, which resulted in a duration of 2 minutes 21 seconds. This number was rounded up to 2 minutes 30 seconds and used as the duration for phase 1. During this pilot study phase 2 ended when the participant had focused on the virtual characters at least once and then when they indicated that they thought they had seen everything there was to see. To determine the final duration of phase 2 for the experiment, the duration that the participant was focused on the GAVC including gaze following behaviour was averaged over all participants. This resulted in 1 minute 6 seconds, which was rounded up to 1 minute 30 seconds for the experiment.

The study was approved by the university's Human Research Ethics Committee (ID: 618) and preregistered using the Open Science Framework (OSF) [32].

4.2. Materials

For the VR Head Mounted Display (HMD) the FOVE 0 [30] was used, which could track eye gaze with an accuracy of less than 1 degree. It had a resolution of 2560x1440 pixels and the eye tracker's latency was 20ms [69]. The virtual environment was created using Unity3D [77]. The thinking aloud audio was recorded using the onboard microphones of a HP Zbook 14 laptop.

4.3. Measurements

Believability

How believable participants thought the virtual characters were was measured using the Believability questionnaire from Gomes et. al. [34]. Participants evaluated to what extent each of the items matched with the virtual character, using a seven-point Likert scale (1: "Totally disagree", 7: "Totally agree") for each item.

Social presence

How socially present participants considered the virtual characters to be was measured using Bailenson's Social Presence questionnaire [5]. Just like for the believability questionnaire, participants evaluated to what extent each of the sentences matched with the virtual character, using a seven-point Likert scale (1: "Totally disagree", 7: "Totally agree").

Gaze behaviour

As an objective measurement, the gaze behaviour of the participant was used. The first behavioural measurement was how often and the total duration that the participant looked at each virtual character during both phases individually. The second behavioural measurement was how often the participant followed the gaze of the virtual characters to objects in the scene during phase 2.

Theory of Mind

One of the ways to measure the application of ToM is called *Triangles*, also known as the Social Perception Task [1, 43]. In this task participants were asked to narrate the movements of inanimate shapes presented as short videos. Responses were recorded, transcribed and scored with more points being awarded when mentalising or attributing thoughts or feelings. This approach was adapted to fit this experiments' specific situation. Participants were asked to think aloud during the experiment; Instructions were based on Boren and Ramey's guidelines [14]. The response from the participants was recorded with the aim of capturing the participants' stream of consciousness. Two coders then scored all moments in which the participant either ascribed reasoning or thinking, or attributed feelings to the virtual characters in 10 second intervals. This ToM scoring form can be found in appendix A.

4.4. Procedure

Participants were first provided with the informed consent form. After signing this form the participants were given a demographic questionnaire. They were then provided with a more thorough explanation of the thinking aloud procedure including an example in the form of a video [72]. After this the eye tracking calibration procedure of the VR HMD was explained and performed. They then moved on to a virtual practice room, where they had some time to get used to VR and practice the thinking aloud procedure as suggested by Boren and Ramey's guidelines on thinking aloud procedures [14]. This room was designed so that users were able to get used to VR and practice thinking aloud without the novelty effect influencing the experimental setup. When they indicated they were ready, the names of the virtual characters were given and they were moved to the virtual waiting room. From this moment on measurements were taken. After 4 minutes the participants were asked to remove the HMD. Then they were asked to fill in the last two questionnaires (believability and social presence), followed by a debriefing. Figure 4.1 shows the experimental setting.

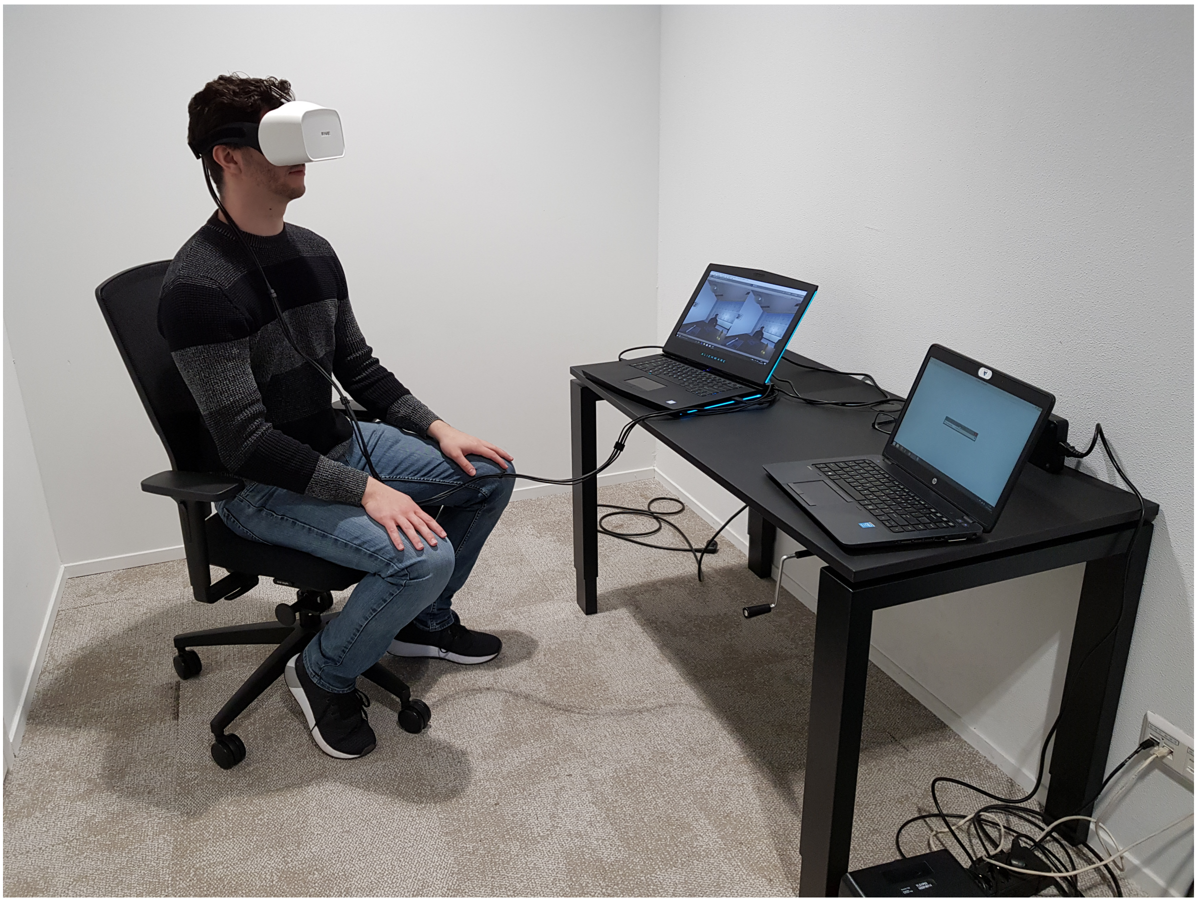


Figure 4.1: A participant during the experiment.

4.5. Participation

A total of 24 participants (12 males, 12 females) participated in the experiment. All participants were recruited on the university campus, either personally or via email. The ages of the participants varied between 20 and 30 (M=24, SD=2.41). The whole experiment procedure took approximately 20 minutes. The experiment was conducted in a small dedicated room. Participation was voluntary and no participants were paid or offered compensation.

4.6. Data preparation and analysis

R, version 3.5.2, was used to conduct statistical analyses. The R markdown scripts are available online [33]. The first step was to prepare the data. For the believability questionnaire this involved a reliability analysis. A Cronbach's alpha of 0.72 for the CVC and 0.54 for the GAVC was found. Which was considered acceptable to use the averages of the items in subsequent analyses. For the social presence questionnaire data preparation involved the scores to be summed up into a single score as the source paper suggests [5]. A positive score indicates that the participant believed the virtual character was conscious and was watching him or her, whereas a negative score indicates that the participant felt the virtual character was just a computerised image. Data preparation for the ToM measure involved comparing the results of the two coders and summing up the individual scores of the 10 second intervals per item per virtual character and separated per phase. Disagreement between coders was solved by joint discussion. In practice coders agreement was 100%. Then, to examine the effects of the virtual characters, either a paired t-test or a Wilcoxon signed-rank test was used depending on the normality of the distribution. Finally, the effect size was calculated for the measurements that reached significance. Cohen's convention [19] was used for this. For parametric tests Cohen's d is used. He suggested that $d = 0.20$ be considered a small effect size, $d = 0.50$ represents a medium effect size and $d = 0.80$ a large effect size. Similarly, for non-parametric tests Pearson's r is used. Following Cohen's convention [19] $r = 0.10$ is a small, $r = 0.30$ is a medium and $r = 0.50$ is a large effect size.

5

Results

Table 5.1, table 5.2 and table 5.3 show an overview of the results.

Believability

The results of the believability questionnaire showed a significant increase ($t = -2.43$, $p = 0.02$, $d = 0.50$). This indicated that participants perceived the GAVC ($M = 4.15$, $SD = 0.77$) as more believable than the control condition ($M = 3.70$, $SD = 1.01$).

Table 5.1

Results for the believability, social presence, gaze behaviour and ToM measurements across phase 1 and 2

Measurement	*z**	*t***	*p*	*d***	*r**
Believability		-2.43	0.02	0.50	
Social presence	-2.26		0.02		0.46
Gaze behaviour					
phase 1, total gaze duration		-0.26	0.80	0.05	
phase 1, number of gazes		1.27	0.22	0.20	
phase 2, total gaze duration		2.37	0.03	0.35	
phase 2, number of gazes		1.64	0.11	0.25	
phase 2, number of gazes to objects	-0.53		0.60		0.11
Theory of Mind					
phase 1, total reasoning	0		1		0
phase 1 total feeling	-1.31		0.19		0.27
phase 2, total reasoning	-0.58		0.56		0.12
phase 2, total feeling	-1		0.32		0.20

**when a non-parametric test has been performed*
*** when a parametric test has been performed*

Social presence

The results of the social presence questionnaire showed a significant increase ($Z = -2.26$, $p = 0.02$, $r = 0.46$). For the GAVC 17 out of 24 scores were positive ($Mdn = 5.00$), while in the control condition 13 out of 24 scores were positive ($Mdn = 2.50$). Indicating that 17% more participants perceived the GAVC to be socially present.

Gaze behaviour

The results of all behavioural measurements except one were not significant. The only significant behavioural measurement ($t = 2.37$, $p = 0.03$, $d = 0.35$), was the total duration looked at the virtual characters (individually) in phase 2. In this phase, participants looked more at the control condition ($M = 20.04$, $SD = 9.76$) than the GAVC ($M = 16.88$, $SD = 8.14$).

Table 5.2

Mean and standard deviation results from measurements analysed using a t-test for both conditions

Measurement	*Mean*		*Standard deviation*	
	CVC	GAVC	CVC	GAVC
Believability	3.70	4.15	1.01	0.77
Gaze behaviour				
phase 1, total gaze duration	31.54	32.08	10.32	11.94
phase 1, number of gazes	32.13	30.17	10.84	8.93
phase 2, total gaze duration	20.04	16.88	9.76	8.14
phase 2, number of gazes	20.54	18.54	7.83	8.27

Theory of Mind

The thinking aloud measurement actually consists of four measurements, two scales with two phases each for a total of four. None of these measurements reached statistical significance. Overall indicators of ToM were low as median values were mainly around 0.

Table 5.3

Median results from measurements analysed using a Wilcoxon signed-rank test for both conditions

Measurement	*Median*	
	CVC	GAVC
Social presence	2.5	5
Gaze behaviour		
phase 2, number of gazes to objects	3	4
Theory of Mind		
phase 1, total reasoning	0	0
phase 1, total feeling	1	0.5
phase 2, total reasoning	0	0
phase 2, total feeling	0	0

6 Discussion & Conclusion

6.1. Discussion

The results supported our second hypothesis. Participants perceived a virtual character as more believable and more socially present when they were gaze aware than when they were not.

The findings for the first hypothesis, which proposed that people look at - and follow the gaze of - a gaze aware virtual character more than they do with a non- gaze aware virtual character, suggest a rejection. In phase 2, participants looked longer at the control than at the gaze aware virtual character, though the number of gazes and the amount of gaze following in this phase were not significantly different between the virtual characters. It is possible that people looked longer at the control in phase 2 because the behaviour of the CVC might have been perceived as stranger than the behaviour of the GAVC. The CVC behaved completely independent from the participant and other social factors which was perceived as shy or asocial by some participants. Negative stimuli have been shown to receive more attention [52] and thus a stranger virtual character could have garnered more attention.

The third and last hypothesis was not supported as the analysis result was inconclusive. This hypothesis predicted that people employ Theory of Mind more often to reason about the mental state of a gaze aware virtual character than they do to reason about the mental state of a non- gaze aware virtual character. Overall indicators of ToM were low. Participants rarely mentioned the virtual characters aloud. Instead they talked about the contents of the room and their possible relation to it. The room contained a lot of objects that were perceived as strange by participants when found in a single room. This was intentional to make sure the participants would actually keep looking around so that the gaze following behaviour could be noticed. However, this could have distracted the participants too much causing them to barely think about the virtual characters consciously.

6.2. Limitations

This study has some limitations as well that people have to be aware of. First, the experiment was held in an artificial virtual environment. The unexpected objects in the virtual waiting room could have distracted the participants too much, potentially causing them to miss important details. Although it is not an unreal setting as real life has plenty of places with strange objects, this could potentially have affected the conscious application of ToM to the virtual characters. Second, this study only manipulated the gaze behaviour joint attention and mutual gaze. The non-verbal behaviour expressed by the virtual characters was not further embedded in other social (non-)verbal behaviours. For example one can think of a head nod after establishing mutual gaze to acknowledge the other person. This particular example was mentioned by several participants who said that after the virtual character failed to acknowledge them they lowered their expectations of the virtual characters. Consistency in behaviour is important as it allows people to predict what will happen when they engage with the virtual character [29, 41, 83]. Third, the participants were recruited on campus and thus all university students or employees. They were also between age 20 and 30. It is therefore not clear to which extent findings generalise to other populations and ages.

6.3. Future work

The work can be extended in several ways. For example, in the current study gaze behaviour was dependent on the participant only, but would a virtual character become even more believable and socially present if the gaze behaviour took all actors present into account?

Moreover, a different way of measuring the application of ToM to virtual characters might be able to show an effect between gaze aware behaviour and ToM. Complex emotions can be distinguished from each other only by paying attention to the eyes [9]. Our study was unable to show an effect on the application of ToM. However, it might be possible to find an effect on the application of ToM to virtual characters by designing a study centered around complex emotions. For example, Baron-Cohen [10] tested the application of ToM by asking participants to match photographs of actors' eyes with cognitive states or complex emotions. In this study a similar measure was used but instead of making people choose between several options (some of which were ambiguous) like in Baron-Cohen's study [10], we had people think aloud and then checked afterwards how often they ascribed cognitive states/emotions. This might not have worked as well, possibly because their conscious thoughts were occupied with other things. However, giving people several options to choose from means the experiment needs to be interrupted and broken up into several parts. As the social setting is very important for this study, interruptions in the flow of the interaction would influence that setting in a negative way. The other measurements were dependent on a continuous flow as well, so it was not possible to take that approach in this study. However, a new study focused solely on the application of ToM would be able to use a measure more similar to Baron-Cohen's measure, which might be more suitable for testing the application of ToM.

Instead of a different measurement, a different setup of the experiment might also be able to show an effect between the application of ToM and gaze aware virtual characters. By increasing the duration of the experiment a novelty effect of the objects in the room might be prevented. Although the downside to that is that the limited social behaviour of the virtual characters would become more obvious, possibly decreasing the participants' perception of the capabilities of the virtual characters. Another option to change the setup of the experiment could be to remove the strange objects from the virtual room and only introduce objects that can be expected in a waiting room. Downside of this approach is that participants would most likely become bored. That would probably decrease their attention and might even introduce negative feelings towards the experiment and the virtual characters.

Furthermore, the use of gaze aware virtual characters could be beneficial for other domains as well, such as for therapeutic systems or in gaming. For Autism Spectrum Disorder (ASD) for example, atypical gaze plays a major role in the development of this disorder [2, 21, 78]. Recent studies have delved into the possibility of using serious games to train social skills including gaze behaviour [18, 35, 65, 71]. It remains, however, to be seen whether eye gaze training (and in particular with gaze aware virtual characters) can improve gaze behaviour in people with ASD or other disorders in which atypical gaze behaviour plays a role. Also, can a gaze aware virtual character increase the impact of therapeutic systems in general by increasing the effectiveness of the interaction by making use of gaze as an extra channel for social signals? Besides therapeutic settings, gaze aware virtual characters can have impact on game applications as well. Virtual characters would be more believable, so that developers might be able to prevent negative judgement of their NPCs (non-player characters) like those of the game Alias for example [76].

6.4. Contributions

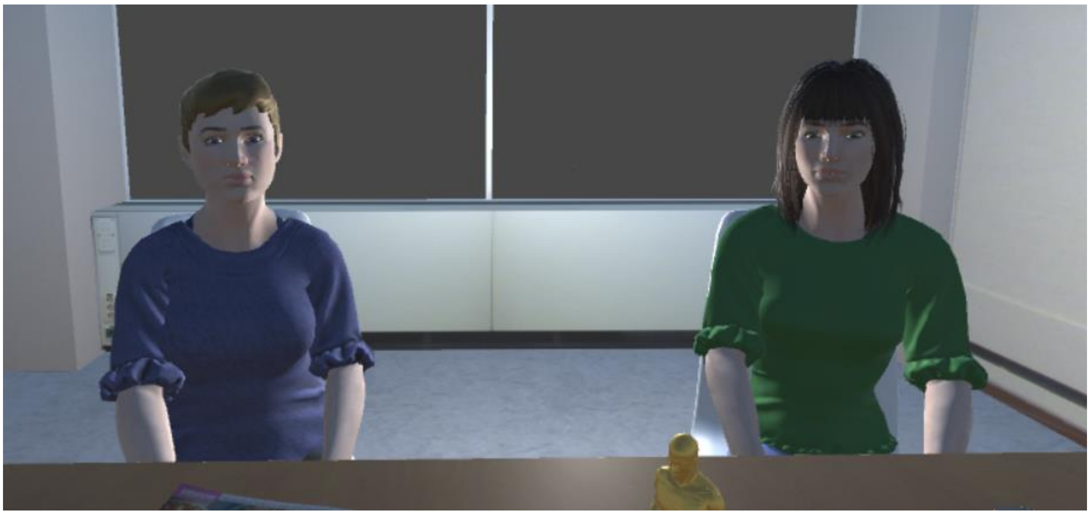
The main scientific contribution of the work is that the perception of believability and social presence of virtual characters can be improved by interacting with a person socially using eye gaze. This work also has a practical contribution for developers of systems with eye-tracking functionality. They could use the gaze behaviour of this study and the ideas for other gaze behaviours to make their virtual characters seem more believable and more socially present in the virtual world. It would help give users the feeling that the virtual characters can impact them socially and that they might be able to impact the virtual characters socially in return. This gives developers some extra handholds to make their virtual characters come to life.

In conclusion this study demonstrates that using eye gaze to interact with people improves the perception of believability and social presence of virtual characters during social interactions.

A

ToM scoring form

Participant number: __________



Alice **Brenda**

If applicable:

1. Participant vaguely started noticing something is different between the avatars at: ________
2. Participant figured out the gaze following of the special avatar at: ______________________

ALICE	0:10	0:20	0:30	0:40	0:50	1:00	1:10	1:20	1:30	1:40	1:50	2:00
3. The participant wonders about or assumes some kind of reasoning or thinking by **Alice**.												
4. The participant attributes a feeling to **Alice**.												
	2:10	2:20	2:30	2:40	2:50	3:00	3:10	3:20	3:30	3:40	3:50	4:00
3. The participant wonders about or assumes some kind of reasoning or thinking by **Alice**.												
4. The participant attributes a feeling to **Alice**.												

BRENDA	0:10	0:20	0:30	0:40	0:50	1:00	1:10	1:20	1:30	1:40	1:50	2:00
3. The participant wonders about or assumes some kind of reasoning or thinking by **Brenda**.												
4. The participant attributes a feeling to **Brenda**.												
	2:10	2:20	2:30	2:40	2:50	3:00	3:10	3:20	3:30	3:40	3:50	4:00
3. The participant wonders about or assumes some kind of reasoning or thinking by **Brenda**.												
4. The participant attributes a feeling to **Brenda**.												

Notes:

Extra information for the coders:

- For statement 1 and 2, if the participant doesn't mention it, write down an "X" instead of a timestamp.
- For statement 1 and 2, take the timestamp of the beginning of the sentence.
- You can assume that the participant correctly identifies which avatar is the special avatar for statement 2.
- For statement 3 and 4, go through the recording in 10sec intervals and for that interval decide whether the statement happened in that interval. If it happened, write down an "X" in that square for the correct avatar.
- If a sentence starts in the first interval and ends in the next interval, write an "X" in both.
- If the participant talks about both avatars (for example: "they are doing this"), write down an "X" for both avatars.
- Statement 3 is about whether the participant mentions that he/she thinks the avatar is reasoning or thinking about something. For example "A keeps looking at the alarm clock, maybe he/she needs to leave soon." or "A seems interested in me." or "A doesn't like me." (because here the participant is assuming that the avatar is thinking about liking the participant). But not "maybe A and B are getting married." (because this is not reasoning by the avatar but about the avatar).
- Statement 4 is about whether the participant mentions that he/she thinks the avatar is feeling something. For example: "A looks sad/happy/emotional/excited" or "A doesn't look happy". But not: "A looks ugly/Slavic/fancy/interested." Negations could also count as long as the participant thinks the avatar has feelings, for example "A doesn't look happy" implies the opposite namely "A looks unhappy", which is a feeling, so it counts.
- The participants were asked to refer to the avatars by their name as much as possible but if they refer to them in a descriptive way you can use the image above the form to determine which one they are talking about.
- If the participant fails to mention which avatar they are talking about and you can't determine this from the context, write down a "?" for both avatars instead.
- If you are doubting, make a decision and write it down normally and then also write the statement, what the participant said and why you are doubting in the notes.

Bibliography

[1] Frances Abell, Frances Happe, and Uta Frith. Do triangles play tricks? attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development*, 15(1):1–16, 2000.

[2] Ashwaq Zaini Amat, Amy Swanson, Amy Weitlauf, Zachary Warren, and Nilanjan Sarkar. Design of an assistive avatar in improving eye gaze perception in children with asd during virtual interaction. In *International Conference on Universal Access in Human-Computer Interaction*, pages 463–474. Springer, 2018.

[3] Michael Argyle and Mark Cook. Gaze and mutual gaze. 1976.

[4] Raúl Arrabales, Jorge Muñoz, Agapito Ledezma, German Gutierrez, and Araceli Sanchis. A machine consciousness approach to the design of human-like bots. In *Believable bots*, pages 171–191. Springer, 2013.

[5] Jeremy N Bailenson, Jim Blascovich, Andrew C Beall, and Jack M Loomis. Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators & Virtual Environments*, 10(6):583–598, 2001.

[6] Jeremy N Bailenson, Jim Blascovich, Andrew C Beall, and Jack M Loomis. Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, 29(7):819–833, 2003.

[7] Jeremy N Bailenson, Andrew C Beall, Jack Loomis, Jim Blascovich, and Matthew Turk. Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. *Human communication research*, 31(4):511–537, 2005.

[8] Simon Baron-Cohen. *Mindblindness: An essay on autism and theory of mind*. MIT press, 1997.

[9] Simon Baron-Cohen, Sally Wheelwright, Jolliffe, and Therese. Is there a“ language of the eyes”? evidence from normal adults, and adults with autism or asperger syndrome. *Visual cognition*, 4(3):311–331, 1997.

[10] Simon Baron-Cohen, Sally Wheelwright, Jacqueline Hill, Yogini Raste, and Ian Plumb. The “reading the mind in the eyes” test revised version: a study with normal adults, and adults with asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2):241–251, 2001.

[11] Joseph Bates et al. The role of emotion in believable agents. *Communications of the ACM*, 37(7):122–125, 1994.

[12] Andrew P Bayliss and Steven P Tipper. Predictive gaze cues and personality judgments: Should eye trust you? *Psychological Science*, 17(6):514–520, 2006.

[13] Frank Biocca, Chad Harms, and Judee K Burgoon. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & virtual environments*, 12(5):456–480, 2003.

[14] Ted Boren and Judith Ramey. Thinking aloud: Reconciling theory and practice. *IEEE transactions on professional communication*, 43(3):261–278, 2000.

[15] Harry Brenton, Marco Gillies, Daniel Ballin, and David Chatting. The uncanny valley: does it exist. In *Proceedings of conference of human computer interaction, workshop on human animated character interaction*. Citeseer, 2005.

[16] Andrew J Calder, Andrew D Lawrence, Jill Keane, Sophie K Scott, Adrian M Owen, Ingrid Christoffels, and Andrew W Young. Reading the mind from eye gaze. *Neuropsychologia*, 40(8):1129–1138, 2002.

[17] Umberto Castiello. Understanding other people's actions: intention and attention. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2):416, 2003.

[18] Leanne Chukoskie, Marissa Westerfield, and Jeanne Townsend. A novel approach to training attention and gaze in asd: A feasibility and efficacy pilot study. *Developmental neurobiology*, 78(5):546–554, 2018.

[19] Jacob Cohen. Statistical power analysis for the behavioral sciences. Technical report, 1988.

[20] Mark Cook. Gaze and mutual gaze in social encounters: How long-and when-we look others "in the eye" is one of the main signals in nonverbal communication. *American Scientist*, 65(3):328–333, 1977.

[21] Matthieu Courgeon, Gilles Rautureau, Jean-Claude Martin, and Ouriel Grynszpan. Joint attention simulation using eye-tracking and virtual humans. *IEEE Transactions on Affective Computing*, 5(3):238–250, 2014.

[22] Daniel C Dennett. *Consciousness explained*. Penguin uk, 1993.

[23] N Doerr. Heavy rain devs have "conquered" the uncanny valley. joystiq, 2007.

[24] Judith Donath. Mediated faces. In *International Conference on Cognitive Technology*, pages 373–390. Springer, 2001.

[25] Quantic Dream. Heavy rain. *Sony Computer Entertainment*, 2010.

[26] Shiri Einav and Bruce M Hood. Tell-tale eyes: children's attribution of gaze aversion as a lying cue. *Developmental Psychology*, 44(6):1655, 2008.

[27] Paul Ekman. *Telling lies: Clues to deceit in the marketplace, politics, and marriage (revised edition)*. WW Norton & Company, 2009.

[28] Nathan J Emery. The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, 24(6):581–604, 2000.

[29] Susan T Fiske and Shelley E Taylor. *Social cognition: From brains to culture*. Sage, 2013.

[30] FOVE. FOVE homepage. `https://www.getfove.com/`, 2019.

[31] David Gamez. Progress in machine consciousness. *Consciousness and cognition*, 17(3): 887–910, 2008.

[32] Garama, Z. (Zilla). OSF preregistration - study on virtual characters with gaze aware social responses in virtual reality. `https://osf.io/8u2kf/`, 2018. [Online; accessed 03-10-2019].

[33] Garama, Z. (Zilla). Data underlying the research of "The role of eye gaze in virtual characters on their perceived believability, social presence and the application of ToM.", 2019. URL `https://data.4tu.nl/repository/uuid:986c1e96-2acf-4e6d-bdde-1c1ad97edf84`.

[34] Paulo Gomes, Ana Paiva, Carlos Martinho, and Arnav Jhala. Metrics for character believability in interactive narrative. In *International Conference on Interactive Digital Storytelling*, pages 223–228. Springer, 2013.

[35] Charline Grossard, Ouriel Grynspan, Sylvie Serret, Anne-Lise Jouen, Kevin Bailly, and David Cohen. Serious games to teach social interactions and emotions to individuals with autism spectrum disorders (asd). *Computers & Education*, 113:195–211, 2017.

[36] Agneta Gulz. Benefits of virtual characters in computer based learning environments: Claims and evidence. *International Journal of Artificial Intelligence in Education*, 14(3, 4): 313–334, 2004.

[37] Nelson G Hanawalt. The role of the upper and lower parts of the face as a basis for judging facial expressions: I. in painting and sculpture. *The Journal of General Psychology*, 27 (2):331–346, 1942.

[38] Marc D Hauser. *The evolution of communication*. MIT press, 1996.

[39] Carrie Heeter. Being there: The subjective experience of presence. *Presence: Teleoperators & Virtual Environments*, 1(2):262–271, 1992.

[40] Bruce M Hood, C Neil Macrae, Victoria Cole-Davies, and Melanie Dias. Eye remember you: The effects of gaze direction on face recognition in children and adults. *Developmental science*, 6(1):67–71, 2003.

[41] Katherine Isbister and Clifford Nass. Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International journal of human-computer studies*, 53(2):251–267, 2000.

[42] Laurent Itti, Nitin Dhavale, and Frederic Pighin. Realistic avatar eye and head animation using a neurobiological model of visual attention. In *Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation VI*, volume 5200, pages 64–79. International Society for Optics and Photonics, 2003.

[43] Michelle R Kandalaft, Nyaz Didehbani, Daniel C Krawczyk, Tandra T Allen, and Sandra B Chapman. Virtual reality social cognition training for young adults with high-functioning autism. *Journal of autism and developmental disorders*, 43(1):34–44, 2013.

[44] Adam Kendon. Some functions of gaze-direction in social interaction. *Acta psychologica*, 26:22–63, 1967.

[45] Chris L Kleinke. Gaze and eye contact: a research review. *Psychological bulletin*, 100 (1):78, 1986.

[46] Tore Knabe. Experiments with Animation for Eyes and Head. `http://tore-knabe.com/experiments-with-animation-for-eyes-and-head/`, 2014.

[47] Tore Knabe. Realistic Eye Movements. `https://assetstore.unity.com/packages/tools/animation/realistic-eye-movements-29168`, 2018.

[48] Stephen RH Langton, Roger J Watt, and Vicki Bruce. Do the eyes have it? cues to the direction of social attention. *Trends in cognitive sciences*, 4(2):50–59, 2000.

[49] Sooha Park Lee, Jeremy B Badler, and Norman I Badler. Eyes alive. In *ACM transactions on graphics (TOG)*, volume 21, pages 637–644. ACM, 2002.

[50] Aaron B Loyall. Believable agents: Building interactive personalities. Technical report, Carnegie-Mellon univ. Pittsburgh PA dept. of Computer Science, 1997.

[51] C Neil Macrae, Bruce M Hood, Alan B Milne, Angela C Rowe, and Malia F Mason. Are you looking at me? eye gaze and person perception. *Psychological science*, 13(5):460–464, 2002.

[52] Linda Marschner, Sebastian Pannasch, Johannes Schulz, and Sven-Thomas Graupner. Social communication with virtual agents: The effects of body and gaze direction on attention and emotional responding in human observers. *International Journal of Psychophysiology*, 97(2):85–92, 2015.

[53] Malia Mason, Bruce Hood, and C Neil Macrae. Look into my eyes: Gaze direction and person memory. *Memory*, 12(5):637–643, 2004.

[54] Malia F Mason, Elizabeth P Tatkow, and C Neil Macrae. The look of love: Gaze shifts and person perception. *Psychological science*, 16(3):236–239, 2005.

[55] Masahiro Mori, Karl F MacDorman, and Norri Kageki. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2):98–100, 2012.

[56] D Morris. *Peoplewatching: The Desmond Morris Guide to Body Language*. Vintage Publishing, 2002.

[57] Clifford Nass and Youngme Moon. Machines and mindlessness: Social responses to computers. *Journal of social issues*, 56(1):81–103, 2000.

[58] Clifford Nass, Jonathan Steuer, and Ellen R Tauber. Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 72–78. ACM, 1994.

[59] Joshua Newn, Fraser Allison, Eduardo Velloso, and Frank Vetere. Looks can be deceiving: Using gaze visualisation to predict and mislead opponents in strategic gameplay. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 261. ACM, 2018.

[60] T. Nummenmaa. *The Language of the Face*, volume 9 of *Jyväskylä studies in education, psychology and social research*. Kustantajat Publ., 1964.

[61] Adam Palanica and Roxane J Itier. Attention capture by direct gaze is robust to context and task demands. *Journal of Nonverbal Behavior*, 36(2):123–134, 2012.

[62] Tomislav Pejsa, Sean Andrist, Michael Gleicher, and Bilge Mutlu. Gaze and attention management for embodied conversational agents. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 5(1):3, 2015.

[63] Kevin A Pelphrey, Jeffrey D Singerman, Truett Allison, and Gregory McCarthy. Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia*, 41(2):156–170, 2003.

[64] Wendy Phillips, Simon Baron-Cohen, and Michael Rutter. The role of eye contact in goal detection: Evidence from normal infants and children with autism or mental handicap. *Development and Psychopathology*, 4(3):375–383, 1992.

[65] Kaśka Porayska-Pomsta, Keith Anderson, Sara Bernardini, Karen Guldberg, Tim Smith, Lila Kossivaki, Scott Hodgins, and Ian Lowe. Building an intelligent, authorable serious game for autistic children and their carers. In *International Conference on Advances in Computer Entertainment Technology*, pages 456–475. Springer, 2013.

[66] Daniel J Povinelli and Todd M Preuss. Theory of mind: evolutionary history of a cognitive specialization. *Trends in neurosciences*, 18(9):418–424, 1995.

[67] David Premack. " does the chimpanzee have a theory of mind?" revisited. *Machiavellian intelligence*, 1988.

[68] Lei Qu, Ning Wang, and W Lewis Johnson. Using learner focus of attention to detect learner motivation factors. In *International Conference on User Modeling*, pages 70–73. Springer, 2005.

[69] Rodrigo. What is the latency of fove eye tracking? `https://support.getfove.com/hc/en-us/articles/115000733714-What-is-the-Latency-of-FOVE-Eye-Tracking-`, 2019.

[70] Brian Scassellati. Imitation and mechanisms of joint attention: A developmental structure for building social skills on a humanoid robot. In *International Workshop on Computation for Metaphors, Analogy, and Agents*, pages 176–195. Springer, 1998.

[71] Sylvie Serret, Stephanie Hun, Galina Iakimova, Jose Lozada, Margarita Anastassova, Andreia Santos, Stephanie Vesperini, and Florence Askenazy. Facing the challenge of teaching emotions to individuals with low-and high-functioning autism using a new serious game: a pilot study. *Molecular autism*, 5(1):37, 2014.

[72] Amber Simpson. Think-aloud method (instructions demo). `https://www.youtube.com/watch?v=-q__eqNevb4`, 2017. [Online; accessed 13-05-2019].

[73] Mel Slater, Andrea Brogni, and Anthony Steed. Physiological responses to breaks in presence: A pilot study. In *Presence 2003: The 6th Annual International Workshop on Presence*, volume 157. Citeseer, 2003.

[74] Harrison Jesse Smith and Michael Neff. Communication behavior in embodied virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 289. ACM, 2018.

[75] Rainer Stiefelhagen, Michael Finke, Jie Yang, and Alex Waibel. From gaze to focus of attention. In *International Conference on Advances in Visual Information Systems*, pages 765–772. Springer, 1999.

[76] Angela Tinwell, Deborah Abdel Nabi, and John P Charlton. Perception of psychopathy and the uncanny valley in virtual characters. *Computers in Human Behavior*, 29(4): 1617–1625, 2013.

[77] Unity Technologies. Unity homepage. `https://unity.com/`, 2019.

[78] Jos Nicolaas Van Der Geest, Chantal Kemner, Marinus N Verbaten, and Herman Van Engeland. Gaze behavior of children with pervasive developmental disorder toward human faces: a fixation time study. *Journal of Child Psychology and Psychiatry*, 43(5):669–678, 2002.

[79] Melodie Vidal, Remi Bismuth, Andreas Bulling, and Hans Gellersen. The royal corgi: Exploring social gaze interaction for immersive gameplay. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 115–124. ACM, 2015.

[80] Hua Wang, Mark Chignell, and Mitsuru Ishizuka. Improving the usability and effectiveness of online learning: How can avatars help? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 49, pages 769–773. SAGE Publications Sage CA: Los Angeles, CA, 2005.

[81] Lawrence Weschler. Why is this man smiling. *Digital animators are closing in on the complex system that makes a face come alive. Wired*, 10, 2002.

[82] Andrew Whiten and RW Byrne. *Natural theories of mind: Evolution, development and simulation of everyday mindreading.* Basil Blackwell Oxford, 1991.

[83] Atef Ben Youssef, Mathieu Chollet, Hazaël Jones, Nicolas Sabouret, Catherine Pelachaud, and Magalie Ochs. Towards a socially adaptive virtual agent. In *International Conference on Intelligent Virtual Agents*, pages 3–16. Springer, 2015.

[84] Jakub Aleksander Zlotowski, Hidenobu Sumioka, Shuichi Nishio, Dylan F Glas, Christoph Bartneck, and Hiroshi Ishiguro. Persistence of the uncanny valley: The influence of repeated interactions and a robot's attitude on its perception. *Frontiers in psychology*, 6:883, 2015.