



DEGREE PROJECT IN COMPUTER SCIENCE AND ENGINEERING,
SECOND CYCLE, 30 CREDITS
STOCKHOLM, SWEDEN 2016

Perception of Trustworthiness and Valence of Emotional Expressions in Virtual Characters

NIKLAS BLOMQVIST

Perception of Trustworthiness and Valence of Emotional Expressions in Virtual Characters

Perception av Pålrltllghet och Valens av Känslouttryck i Virtuella Karaktärer

Niklas Blomqvist
nblomqvi@kth.se

DA225X Degree Project in Computer Science and Communication, Second Cycle 30 credits
Degree Programme in Computer Science and Engineering, 300 credits
Master's Programme in Human-Computer Interaction, 120 credits

KTH Royal Institute of Technology
School of Computer Science and Communication (CSC)

Supervisor: Christopher Peters

Examiner: Tino Weinkauff

2016-07-05

KTH CSC
SE-100 44 Stockholm, Sweden
URL: www.kth.se/csc

SAMMANFATTNING

Kunskap om hur man designar pålitliga virtuella karaktärer är viktigt när dessa blir mer och mer vanliga interaktionspartners. Den här studien tar en närmare titt på den, via tidigare forskning, föreslagna relation mellan valens och pålitlighet genom att konstruera virtuella karaktärer med olika non-verbala beteenden och låta deltagare betygsätta dessa i en förstudie. En andra fråga om hur perceptionen av pålitlighet grundas i virtuella karaktärer undersöks genom att låta deltagare spela ett trust game med virtuella karaktärer i naturlig storlek på en stor 4k-skärm. Resultaten antyder på att valens inte är en faktor som påverkar pålitlighet och att positiv valens tillsammans med ögonkontakt inte är tillräckligt för att skapa en klart pålitlig virtuell karaktär. Resultaten pekar även på att perceptionen av pålitlighet inte bara är baserad på en virtuell karaktärs tidigare beslut i en längre interaktion utan även på dess icke-verbala beteende. Studiens resultat kan hjälpa till med att konstruera virtuella karaktärer för olika scenarion, speciellt när målet är att göra dom så pålitliga som möjligt. Studien ger också insikt inom verktyg och mjukvaror som kan användas när man skapar virtuella karaktärer och scenarier för förtroende.

ABSTRACT

Knowledge on how to design trustworthy virtual characters are of importance when these are becoming more and more common interaction partners. In this study, a closer look at the suggested relationship from previous research between valence and trustworthiness is investigated by constructing virtual characters with different non-verbal behaviours and letting participants rate them in a pre-study. A second question of how perception of trustworthiness is based for virtual characters is investigated by letting participants play a trust game with life-sized virtual characters on a big 4k-screen. Results indicated that valence is not necessarily a factor influencing trustworthiness and that positive valence together with mutual gaze is not enough to provide a clearly trustworthy virtual character. Results also indicated that perception of trustworthiness is not based solely on a virtual character's previous decisions of trust in a longer interaction but also on its non-verbal behaviour. The outcome of this study will help when constructing virtual characters in different scenarios, especially when the goal is to make them as trustworthy as possible. The study also gives insight into tools and software that can be used when creating virtual characters and setting up scenarios of trust.

Perception of Trustworthiness and Valence of Emotional Expressions in Virtual Characters

Niklas Blomqvist
Master Thesis
KTH Royal Institute of Technology
Stockholm, Sweden
nblomqvi@kth.se

ABSTRACT

Knowledge on how to design trustworthy virtual characters are of importance when these are becoming more and more common interaction partners. In this study, a closer look at the suggested relationship from previous research between valence and trustworthiness is investigated by constructing virtual characters with different non-verbal behaviours and letting participants rate them in a pre-study. A second question of how perception of trustworthiness is based for virtual characters is investigated by letting participants play a trust game with life-sized virtual characters on a big 4k-screen. Results indicated that valence is not necessarily a factor influencing trustworthiness and that positive valence together with mutual gaze is not enough to provide a clearly trustworthy virtual character. Results also indicated that perception of trustworthiness is not based solely on a virtual character's previous decisions of trust in a longer interaction but also on its non-verbal behaviour. The outcome of this study will help when constructing virtual characters in different scenarios, especially when the goal is to make them as trustworthy as possible. The study also gives insight into tools and software that can be used when creating virtual characters and setting up scenarios of trust.

Keywords

Virtual Characters; Trust; Trustworthiness; Valence; Emotion; Facial Expression; Gaze; Body Behaviour; Perception; Trust Game; Prosocial Behaviour

1. INTRODUCTION

There is an increasing need for designing virtual characters (VCs). They are most commonly used in games and animation but are also being used in education, training, healthcare and electronic commerce (e-commerce). Research have shown that virtual characters can improve rapport and learning in education [11], collaboration [16] and therapy [21]. In e-commerce, VCs can be used as intermediates between websites and users, adding multi-modality channels and interaction [6][17]. Designing VCs are also of importance to human behaviour and perception research. Experiments can be set up in a lab environment where the interaction and appearance of the VC are manipulated with a set of variables rather than having a human mimic them. The experimental design is also more easily reproduced using VC stimuli instead of human.

We make decisions of trust every day, from purchasing something from the store or relying on that the store sells



Figure 1: The 3 non-verbal behaviours used in the trust game scenario in the main study. Hypothesized untrustworthy (Insecure_1) to the left, hypothesized neutral (Neutral_1) in the middle and hypothesized trustworthy (Smiling_2) to the right

genuine products and not counterfeits of the brand to electing politicians or trusting our monetary system. Do we borrow our phone to a stranger, do we trust the friend to babysit? We cope with these uncertainties every day by relying on trust, which is the basis for all social interaction, even between humans and VCs. [1] There are many definitions of trust and for this report the following definition is used: "Trust is the willingness to accept vulnerability based upon positive expectations about another's behaviour". [18] There are also different types of trust, specifically 3 types: Interpersonal Trust, System Trust or Impersonal Trust and lastly, Dispositional Trust. Interpersonal Trust is context specific, meaning that person A might trust another person B, who is a mechanic, with servicing the car but not to babysit. Impersonal Trust refers to trust based on perceived properties or reliance in a system in which that trust exists. An example is our monetary system. Dispositional Trust, sometimes referred to as basic trust, is independent any party or context and is a sense of trust in oneself and the world. [12]

Trusting VCs are equally as important for interaction as it is for human interaction and this study aims to investigate how to construct trustworthy and untrustworthy VCs but also to investigate how the perception and impression of the VCs are influenced when context, such as the decisions made in a trust game (see section 3.4.4), is added.

Previous research ties non-verbal behaviour such as smiling and mutual gaze to both trustworthiness and positive

Table 1: The personalities of the VCs are made up of a non-verbal behaviour and a decision algorithm

Definition	Non-Verbal Behaviour	Decision Algorithm	Label
Congruent Untrustworthy	Hypothesized Untrustworthy	Untrustworthy	CU
Incongruent Trustworthy	Hypothesized Untrustworthy	Trustworthy	IT
Congruent Neutral	Hypothesized Neutral	Neutral	CN
Incongruent Untrustworthy	Hypothesized Trustworthy	Untrustworthy	IU
Congruent Trustworthy	Hypothesized Trustworthy	Trustworthy	CT

valence, and anger and averted gaze to untrustworthiness and negative valence.

Two pilot studies (3x2 participants), one pre-study (21 participants) and a main study (21 participants) is conducted. In the main study, participants play a trust game with 5 different VCs who are constructed with a non-verbal behaviour (consisting of facial expressions, gaze behaviour and body behaviour) and a decision algorithm (controlling how much the VCs will give back in the trust game scenario) (see table 1). In the trust game, participants are given an amount of credits each round and given a decision to either keep the credits or give any to the VC. Any amount given will be tripled and the VC will then decide to give something back. The goal is to earn as many credits as possible.

The VCs are given a random appearance of equal trustworthiness and valence (annotated in the pre-study) so that the appearance should have no impact on the results.

1.1 Report outline

This report will begin with a relevant research section (see section 2) where research regarding trustworthiness and VCs is discussed. The VCs in this study are created based on these findings. After that, the method section (see section 3) where two pilot studies, a pre-study and a main study is described. In this section, the specifics on how the VCs are created and what differentiates them is described. The errors caught in the pilot studies and the results of the pre-study and main study are then described in the results (see section 4) followed by a discussion (see section 5). In the discussion, critical discussion (see section 5.6) is considered and at the end of the report is the conclusion (see section 6) and future research (see section 6.1) followed by acknowledgements (see section 7).

2. RELATED RESEARCH

This section will start off by describing research in relation to trustworthiness and VCs. After that, the purpose statement is presented and a research question together with two hypotheses are introduced (see section 2.7).

2.1 Trustworthiness in the trust game

Trustworthiness plays an important role in a risky decision-making game such as the trust game. Counterparts who are perceived as more trustworthy receives more money in the trust game. [22]

2.1.1 Prosociality

Prosocial behaviour is social behaviour that benefits others or society as a whole. The capacity to detect prosocial

tendencies in others allows individuals to strategically enter into cooperative, mutually beneficial relations and avoid the cost of being exploited. The trust game is one such scenario where participants benefit from detecting prosocial tendencies - trust tendencies to determine whether the counterpart will cooperate in a prosocial manner or not. Mutual gaze and smiling are two non-verbal behaviours that are signalling a prosocial behaviour. [8]

2.2 Facial Expressions

Research have shown that facial expressions relate to trustworthiness. In a recent study they showed that face trustworthiness modulates the intensity of perceived emotions: trustworthy faces that expressed happiness were perceived happier than untrustworthy faces and untrustworthy faces that expressed anger were perceived as angrier than trustworthy faces expressing the same emotion. [15] Another study found that in situations with financial stakes, smiling counterparts were rated as more trustworthy, likeable, attractive and expected to cooperate than non-smiling (neutral) counterparts. They also found that fake smiling was rated higher than no smiling at all. [10]

There have been similar results in other studies. For example, smiling can act as a signal, intended to introduce cooperative behaviour and that counterparts in a trust game scenario that smiles are trusted more than non-smiling counterparts. Males have also been found to be more cooperative, especially towards images of female counterparts and that females are least likely to cooperate with other females. [19]

2.3 Appearance

Static facial features such as the facial width-to-height ratio in males have also shown to have a direct correlation with trust. Males with a higher facial width-to-height ratio are less trustworthy and are also perceived as less trustworthy. [20] Moreover, attractiveness have been shown to influence participants actions in the trust game. An attractive counterpart is viewed as more trustworthy and is trusted more than less attractive counterparts. As a result, they earn more in the first stage of a trust game but due to this, they have a higher expectation on them. If this is not lived up to, they are penalised for it, receiving less in the next stage of the trust game. [24]

Dynamic facial features such as smiling and eyebrow movement are more connected to trustworthiness and that static features like facial width-to-height ratio is more connected to physical ability. [7]

2.4 Valence

A term used to describe an emotion as positive or negative is valence. Happiness and gratitude are two emotions with positive valence that have been shown to increase trust, while anger, an emotion with negative valence have been shown to decrease it. [5]

Functional imaging studies have suggested that decisions about trustworthiness involve brain structures (such as the amygdala) that process emotions which could indicate a correlation between the two. [25]

Other research have found that faces are evaluated on two fundamental dimensions: valence and dominance, where valence is an over-generalization of facial cues signalling whether to approach or avoid a person. They show that participants'

judgements of attractiveness were highly correlated to both trustworthiness and valence judgements. [14]

2.5 Gaze and eye behaviour

Other factors that have been shown to influence trustworthiness is gaze and that there is no difference in perception of gaze for humans and VCs. [2] Earlier work with embodied conversational agents have shown that gaze and eye movement are not only essential for non-verbal communication, but also more crucial than expressions. [4] Mutual gaze together with neutral or happy faces increases perceived trustworthiness and mutual gaze together with angry faces decreases it. [23]

Another study found that a higher percentage of mutual gaze is perceived as more trustworthy than averted gaze and that the effect can hold for different facial expressions and scene contexts. They also found that this was true when participants viewed stimuli with a 5 second duration. [13]

2.6 Body behaviour

Body behaviour is another variable that have been shown to influence the perception of VCs. For example, face scratching and arms crossing conveys a defensive and dishonest message. [17] A study investigating how different body postures in the three cultures: American, Sri Lankan and Japanese, found that while there are similarities between cultures in how they convey, recognize and attribute emotional meaning to posture, there are also differences. [9]

2.7 Purpose statement

With the increasing areas of use for VCs, it is important to know how they are perceived in terms of trustworthiness based on their non-verbal behaviour, and also how their decisions influence that perception when focus is moved from the first impression to added context through an interaction.

2.7.1 Research Question

"How does the non-verbal behaviour and decisions of a virtual character influence its perceived trustworthiness?"

2.7.2 Hypotheses

Previous research suggests a relationship between smiling and trustworthiness. [10][19] Research also suggests a relationship between positive valence (where smiling is included) and trustworthiness. [5][14][15] Gaze also seem to be a factor influencing trustworthiness and that mutual gaze increases trustworthiness together with neutral or happy faces. [23][13] Moreover, little research on the relationship between context and non-verbal behaviour have been found. Based on this, the following two hypotheses are defined:

1. A virtual character with a positive valence together with mutual gaze will be perceived as trustworthy
2. The perceived trustworthiness in a virtual character will be solely based on its previous decisions of trust in a longer interaction

Hypothesis 1 will be investigated by constructing VCs based on trustworthy non-verbal behaviours from research and then rated in both a pre-study and a main study in terms of trustworthiness and valence. Hypothesis 2 will be investigated in the main study, where participants play a longer trust game with 5 different VCs. (see table 1)

3. METHOD

A pre-study with 21 participants was conducted to annotate appearances and non-verbal behaviours in terms of trustworthiness and valence. Some of these were chosen and put together for the main study where participants played a trust game (see section 3.4.4) with 5 VCs, which is a commonly used method to quantitatively measure trust. In the main study, participants played 3 rounds with each of the 5 VCs, 3 times, for a total of 9 rounds with each VC, to study the scenario in a longer and repeating interaction. Before both pre-study and main study, a pilot study each was conducted with 3 participants.

3.1 General stimuli

3.1.1 Virtual characters

Virtual characters (VCs) were created using Unity¹ and were based on the MCS Female² model. Since previous research have implied a difference in perception for females and males in the trust game, this study focused on female counterparts. [19]

Two assets^{3,4} were used to add more diversity to the appearances. To manipulate the gaze behaviour, the asset Realistic Eye Movements⁵ was used. Body behaviour was created by manipulated animations from the Taichi Character Pack.⁶ Stimuli was presented on a white background.

3.1.2 Software

To record the stimuli created in Unity, ZD Soft Screen Recorder v9.1 was used⁷. Stimuli were recorded in 30 fps for 10 second durations and saved as .avi-files. To expose the stimuli to participants and record their answers, Tobii Pro Studio⁸ was used in the pre-study. In the main study, exposure and recording of answers were executed in a built Unity file. The data collected was analysed using Microsoft Excel⁹ and SPSS¹⁰.

3.2 Pilot studies

Before the pre-study, a pilot study was held with 3 participants to catch errors in the design. This was also done before the main study with 3 participants. Participants did a think-aloud evaluation where they were performing the same tasks as the upcoming participants would for these studies but they would also think aloud, i.e, speak about what they thought when doing so. Errors caught are discussed in the results (see section 4.1)

¹<https://unity3d.com/>

²<https://www.assetstore.unity3d.com/en/#!/content/45807>

³<https://www.assetstore.unity3d.com/en/#!/content/45811>

⁴<https://www.assetstore.unity3d.com/en/#!/content/51541>

⁵<https://www.assetstore.unity3d.com/en/#!/content/29168>

⁶<https://www.assetstore.unity3d.com/en/#!/content/15667>

⁷<https://www.zdsoft.com/>

⁸<http://www.tobii.com/product-listing/tobii-pro-studio/>

⁹https://en.wikipedia.org/wiki/Microsoft_Excel

¹⁰<https://en.wikipedia.org/wiki/SPSS>



Figure 2: To the left is appearance F and to the right is the non-verbal behaviour Angry_1 on appearance A

3.3 Pre-study

The goal of the pre-study was to annotate different appearances in terms of trustworthiness and valence and to see if there was any clear relationship between positive valence and mutual gaze, and trustworthiness. Five appearances and 3 non-verbal behaviours were needed for the the main study.

3.3.1 Participants

Twenty-one students (20 of which between the ages 20-30, 16 men and 5 women) participated in the pre-study. Participants were gathered by holding a presentation in the course Computer Graphics and Interaction at the Royal Institute of Technology, which was followed by inviting course attendants to the pre-study for course credits.

3.3.2 Stimuli

Twenty-three stimuli was shown for 10 second durations in a random order using a Latin square algorithm¹¹, divided into two blocks (23 samples each). The stimuli consisted of video clips of virtual characters in 4 different ways. The stimuli investigated in this study was of 7 appearances and 7 non-verbal behaviours (see fig. 2).

The non-verbal behaviours was constructed using facial expressions, body behaviour, mutual gaze chance and how often the VC's eyes do micro and macro saccades (more eye saccades gives the VC a nervous look). The 7 non-verbal behaviours annotated in the pre-study can be found in table 2.

3.3.3 Procedure

One or two participants performed the pre-study at a time and they were firstly greeted and asked to take a seat (see fig. 3). They were then instructed to read the debrief and sign the consent form. They got to familiarize themselves with the pre-study by testing it on two random virtual characters. During this period, they were allowed to ask questions. After the pre-study was finished they were asked to fill in a demographics form and asked if they wanted to sign up for the main study.

¹¹https://en.wikipedia.org/wiki/Latin_square

Table 2: The non-verbal behaviours used as stimuli in the pre-study

Name	Facial Expression	Body Behaviour	Mutual Gaze	Saccade
Angry_1	Angry	Crossed Arms	80% chance	Calm
Angry_2	Insecure	Crossed Arms	30% chance	Calm
Insecure_1	Insecure	Hand scratching neck	30% chance	Nervous
Neutral_1	Neutral	Neutral	30% chance	Calm
Smiling_1	Smiling	Neutral	80% chance	Calm
Smiling_2	Smiling (increased)	Neutral	80% chance	Calm
Smiling_3	Smiling (very increased)	Neutral	80% chance	Calm

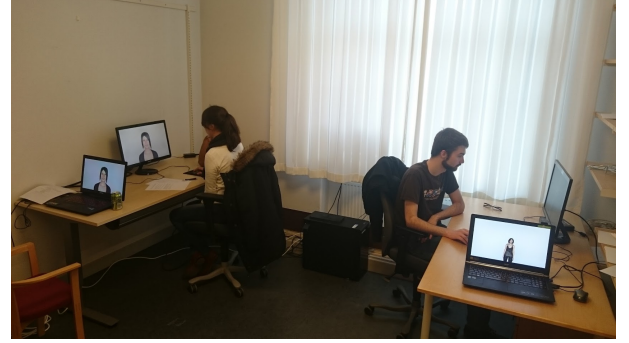


Figure 3: Setup of the pre-study

Participants were seated approximately 60 cm from the computer screen (LED, 24", 1920x1080 resolution) where stimuli were presented for 10 seconds, followed by two questions of trustworthiness and valence. These were rated in a 5 point Likert scale from 1 to 5 (valence: 1=very negative, 3=neutral, 5=very positive; trustworthiness: 1=very untrustworthy, 3=neutral, 5=very trustworthy). The pre-study took about 30 minutes to complete.

Within-subjects ANOVAs were conducted to spot differences between appearances and between non-verbal behaviours in terms of trustworthiness and valence. Post-hoc tests (with Bonferroni adjustment for multiple comparisons) were used when these differences were found to check which groups were significantly different from each other.

3.4 Main study

In the main study, participants played a trust game with 5 different VCs with unique personalities (see table 3). The goal was to investigate how the VCs are perceived from initial impression to when interaction has been added.

3.4.1 Participants

Twenty-one students (19 of which between the ages 20-30, 16 men and 5 women) participated in the main study. Fifteen participants were invited from the pre-study.

3.4.2 Stimuli

The 5 VCs used in the trust game were rated in terms of valence in the pre-study. They had different chances of mutual gaze (which lasted between 2 to 6 seconds) and decision algorithms deciding how much to give back (see table 3 for detailed information). The VCs were randomly mapped to one of 5 appearance for each participant, annotated in the pre-study and matched for equal trustworthiness and valence.

Table 3: VCs used in the trust game had been annotated in the pre-study in terms of valence, they had a specific gaze behaviour and a decision algorithm

Name	Valence	Mutual gaze	Decision algorithm	Label
Congruent Untrustworthy	Negative	30% chance	20-40% back	CU
Incongruent Trustworthy	Negative	30% chance	160-180% back	IT
Congruent Neutral	Neutral	30% chance	100-120% back	CN
Incongruent Untrustworthy	Positive	80% chance	20-40% back	IU
Congruent Trustworthy	Positive	80% chance	160-180% back	CT

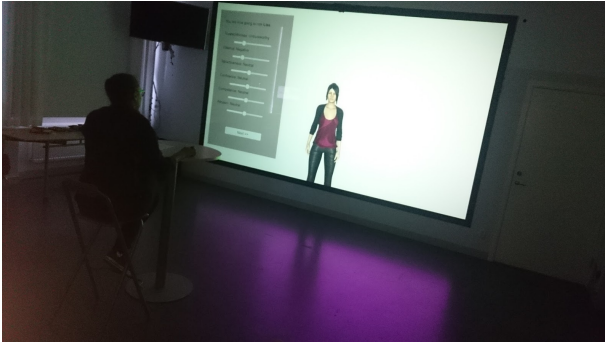


Figure 4: Setup of the main study

3.4.3 Procedure

The main study was held in the Visualization Studio at the Royal Institute of Technology. Participants were greeted and instructed to read the debrief and sign the consent form. They then got to try out the trust game with a random VC for 3 rounds. During this period, they were allowed to ask questions. After the study, participants were asked to fill in a demographics form.

Participants were seated in the middle of the room approximately 3 meters from a 4k-screen (4x2.4 meters, 4096x2160 resolution) on the wall (see fig. 4). The VCs were positioned as they were standing on the ground and their size were that of a real human being to add realism to the scenario. Participants used a mouse to interact in the trust game by either clicking on buttons or dragging sliders.

Prior to the trust game, participants were informed that their results would impact what type of prize they would get. They did not know beforehand what the prizes were and after the study they got to choose between candy (average cost of 1.80\$) or if they ended up with 5500 credits or more in the trust game, they could also choose a Trisslott (3.60\$). Before and after playing the trust game with each VC, participants rated them in terms of trustworthiness, valence, attractiveness, confidence, competence and altruism. This was done using a 5 point Likert scale from 1-5. The main study took around 20-30 minutes per participant.

Within-subjects ANOVAs were conducted to spot differences. Post-hoc tests (with Bonferroni adjustment for multiple comparisons) were used when these differences were found to check which groups were significantly different from each other.

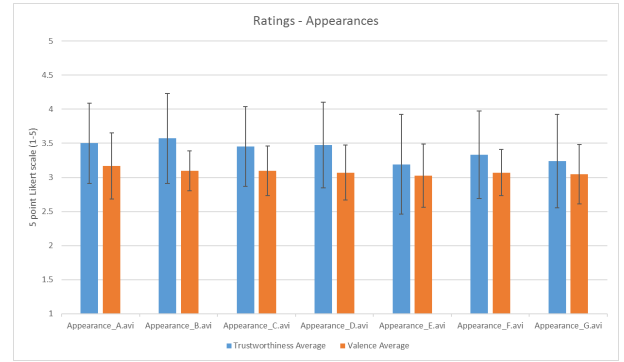


Figure 5: The appearances rated for trustworthiness and valence in the pre-study

3.4.4 Trust game

The trust game used in the main study is an instance of the investment game [3] and is commonly used for quantitatively measuring trust and trustworthiness. [22][24][10][20] Participants were endowed with 100 credits per round and given the choice to either keep all the credits or give any amount to the VC. Any amount given is tripled and then the VC decides whether to give any amount back. Participants played 3 rounds with each VC in a random order (using the fisher-yates shuffle¹²) divided into 3 blocks. This resulted in a total of 45 rounds where 9 (3x3) rounds were played with each VC.

4. RESULTS

In this section, all the results from the pre-study and the main study will be reported. First the errors caught in the pilot studies will be discussed (see section 4.1). Then, the results from the pre-study will be presented (see section 4.2) followed by the results from the main study (see section 4.3). In the main study section, the stimuli chosen will be discussed briefly (see section 4.3.1) followed by a comparison of results between pre-study and main study (see section 4.3.2). After this, the participants' ratings before and after the trust game is reported (see section 4.3.3) and lastly, the betting process (see section 4.3.4) is presented.

4.1 Pilot studies

In the pilot study before the pre-study, participants noticed a few errors and unclear statements in the debrief. They also noted the need to be able to test the study beforehand on a few VCs. This was added before the pre-study.

In the pilot study before the main study, participants had a few thoughts about the descriptive text within the trust game. These were changed before the main study. The neutral decision algorithm was also changed after feedback here.

The initial neutral algorithm was that VCs would give back between 95-115% of any given amount, but participants thought the VCs giving back a negative amount were clearly untrustworthy which was not the intention.

¹²https://en.wikipedia.org/wiki/Fisher%E2%80%93Yates_shuffle

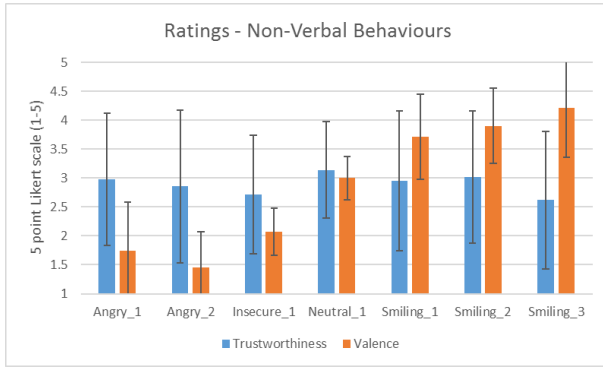


Figure 6: The non-verbal behaviours rated for trustworthiness and valence in the pre-study

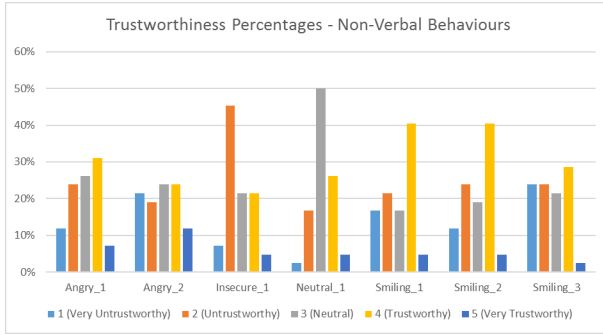


Figure 7: The pre-study ratings of trustworthiness for the non-verbal behaviours in percentage

4.2 Pre-study

The 7 appearances used as stimuli (see fig. 5) in the pre-study were rated, on average, as neutral in terms of valence ($M=3.08$, $SD=0.40$) and slightly trustworthy in terms of trustworthiness ($M=3.39$, $SD=0.65$). No significant difference could be found between trustworthiness ($F(6,287) = 1.953$, $p<0.05$, $p<0.001$) or valence ($F(6,287)=0.521$, $p<0.05$, $p<0.001$).

The non-verbal behaviours (see fig. 6) annotated in the pre-study used the appearance of A (see fig. 2). No significant difference could be found between ratings of trustworthiness ($F(6,287)=1.051$, $p<0.05$, $p<0.001$) but a significant difference between ratings of valence were found ($F(6,287)=115.96$, $p<0.05$, $p<0.001$). Neutral_1 was rated the most neutral ($M=3.00$, $SD=0.38$) and post-hoc tests showed that it was rated as significantly different than all other non-verbal behaviours. Post-hoc tests did also show that Angry_1, Angry_2 and Insecure_1 were all rated as significantly more negative than any other non-verbal behaviour with Angry_2 rated the most negative ($M=1.45$, $SD=0.62$). Moreover, all 3 smiling non-verbal behaviours were rated as significantly more positive than any other non-verbal behaviour and Smiling_3 was rated as most positive ($M=4.21$, $SD=0.86$).

4.3 Main study

4.3.1 Stimuli chosen

The setup of the trust game in the main study required

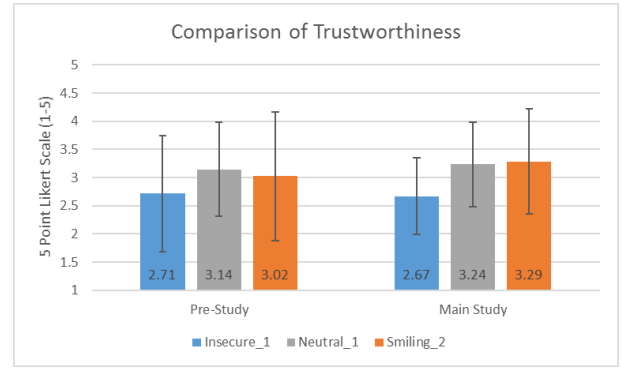


Figure 8: Comparison of the ratings of trustworthiness in the pre-study and the main study

5 appearances perceived as roughly the same in terms of valence and trustworthiness and 3 different non-verbal behaviours, one untrustworthy, one neutral and one trustworthy. The untrustworthy and trustworthy non-verbal behaviours would preferably have the same intensity.

The appearances chosen were A,B,C,D and F. No significant difference in terms of valence or trustworthiness could be found for any of the appearances but appearances E and G were dismissed because they had the highest ratings of untrustworthiness (19% and 12%) in the appearance pool.

No significant difference could be found for the non-verbal behaviours in terms of trustworthiness but looking at the rating percentages (see fig. 7), a tendency could be found where Neutral_1 was rated as more neutral (50%) than any other non-verbal behaviour, Insecure_1 (45%) as more negative and Smiling_1 and Smiling_2 as more positive (40%).

The 3 non-verbal behaviours chosen were Insecure_1, Neutral_1 and Smiling_2. These were annotated in terms of valence as negative ($M=2.07$, $SD=0.40$), neutral ($M=3.00$, $SD=0.38$) and positive ($M=3.90$, $SD=0.65$).

4.3.2 Comparison of ratings

No significant change (see fig. 8) could be found in ratings of trustworthiness for the three non-verbal behaviours chosen (Insecure_1, Neutral_1 and Smiling_2) in the pre-study and the main study ($F(5, 225)=3.111$, $p<0.05$, $p<0.001$).

4.3.3 Ratings

In the following sections, the VCs will be called by their labels (Congruent Untrustworthy=CU, Incongruent Trustworthy=IT, Congruent Neutral=CN, Incongruent Untrustworthy=IU and Congruent Trustworthy=CT) (see table 3).

Participants initial ratings of trustworthiness, before playing the trust game with the VCs, were similar to each other but a significant difference could be found ($F(4,100)=4.474$, $p<0.05$). Post-hoc tests showed that CT ($M=3.52$, $SD=0.91$) was rated as significantly more trustworthy than CU ($M = 2.62$, $SD=0.72$) and IT ($M=2.71$, $SD=0.63$). No significant difference were found between the other VCs.

A significant difference between ratings of trustworthiness, after participants had played the trust game with them, could be found ($F(4,100)=56.47$, $p<0.05$) for the VCs having different decision algorithms. Post-hoc tests showed that the two VCs making trustworthy decisions (giving back 160-180% of any given amount) were not significantly dif-

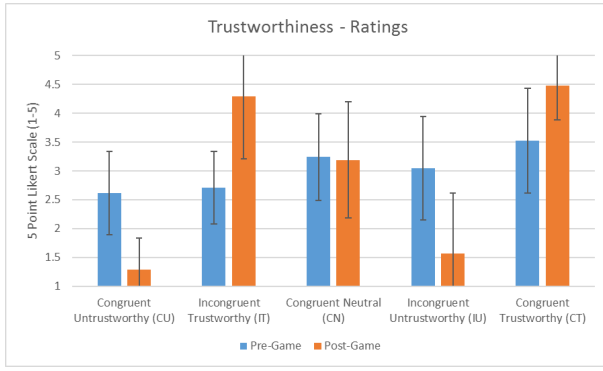


Figure 9: The 5 VCs rated before and after the trust game in terms of trustworthiness

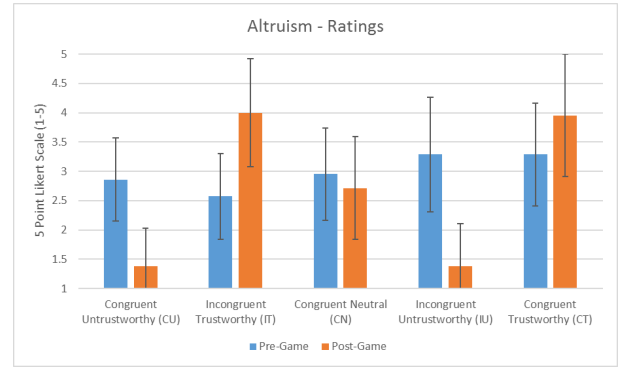


Figure 11: The 5 VCs rated before and after the trust game in terms of altruism

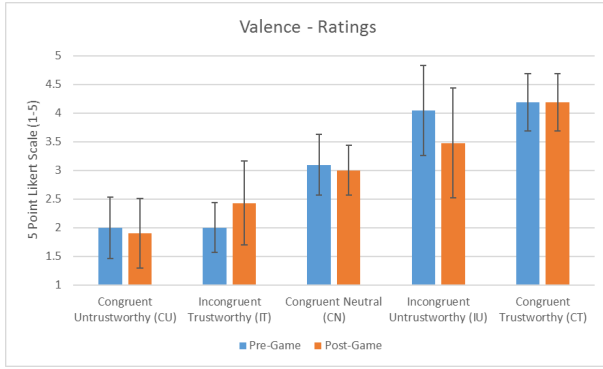


Figure 10: The 5 VCs rated before and after the trust game in terms of valence

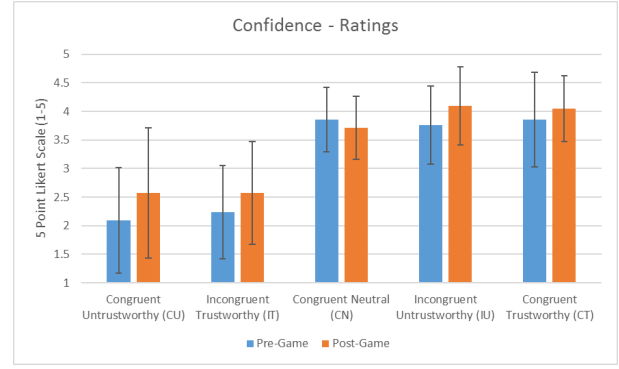


Figure 12: The 5 VCs rated before and after the trust game in terms of confidence

ferent from each other but significantly different from the other VCs and rated between trustworthy and very trustworthy ($M=4.29$, $SD=1.08$ and $M=4.48$, $SD=0.59$). The two VCs making untrustworthy decisions (giving back 20-40% of any given amount) were not significantly different from each other but significantly different from the other VCs and rated between untrustworthy and very untrustworthy ($M=1.57$, $SD=1.05$ and $M=1.29$, $SD=0.55$). The VC making neutral decisions (giving back 100-120% of any given amount) was rated significantly different than all other VCs and rated neutral before and after the trust game ($M=3.24$, $SD=0.75$ and $M=3.19$, $SD=1.01$).

No significant difference between pre or post valence ratings (before and after the trust game) could be found ($F(3,80) = 31.355$, $p<0.05$, $p<0.001$). A small change could be seen for the VCs with incongruent personalities (IT and IU) (see fig. 10). The valence for IT increased after the game ($M=2.00$, $SD=0.43$ to $M=2.43$, $SD=0.73$) and decreased for IU ($M=4.05$, $SD=0.79$ to $M=3.48$, $SD=0.96$).

The ratings of altruism (selfish to unselfish) followed the pattern of trustworthiness ratings (see fig. 11). The VCs with a hypothesized neutral non-verbal behaviour (CN) and hypothesized trustworthy non-verbal behaviour (IU and CT) were perceived as confident both before and after the game and the VCs with a hypothesized untrustworthy non-verbal behaviour (CU and IT) were perceived as nervous (see fig. 12).

No significant difference between pre and post game ratings could be found for attractiveness ($F(9,200) = 7.953$,

$p<0.05$) or competence ($F(9,200)=7.963$, $p<0.05$). The graphs for attractiveness (see fig. 13) and competence (see fig. 14) seems to follow a similar pattern, where the VC consisting of the non-verbal behaviour Neutral_1 is rated most attractive and most competent prior to the trust game followed by the VCs consisting of Smiling_2. The VCs consisting of the non-verbal behaviour Insecure_1 were rated as least attractive and competent prior to the trust game.

4.3.4 Bet process

The bet progression for VCs with the same decision algorithm followed a similar pattern (see fig. 15). The VCs with a hypothesized trustworthy non-verbal behaviour (IU and CT) had similar initial investments ($M=46.29$, $SD=26.49$ and $M=43.95$, $SD=29.31$). The VCs with a hypothesized untrustworthy non-verbal behaviour (CU and IT) did also have similar initial investments ($M=36.90$, $SD=17.47$ and $M=33.95$, $SD=19.59$). The VC with a hypothesized neutral non-verbal behaviour (CN) had a similar initial investment ($M=43.19$, $SD=26.54$) as the VCs with a hypothesized trustworthy non-verbal behaviour (IU and CT).

The VC with the congruent trustworthy personality (CT) received, on average, the most credits ($M=79.29$, $SD=28.55$) as well as the highest last (bet 9) investment ($M=97.90$, $SD=5.28$). The VC with the congruent untrustworthy personality (CU) received, on average the least credits ($M=11.97$, $SD=21.90$) and the lowest last (bet 9) investment ($M=0.52$, $SD=1.79$) (see fig. 16).

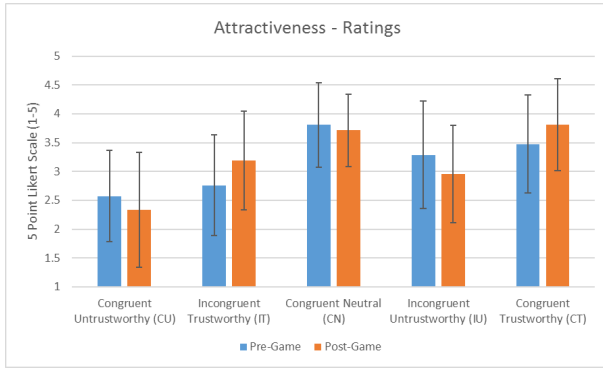


Figure 13: The 5 VCs rated before and after the trust game in terms of attractiveness

5. DISCUSSION

5.1 Valence and trustworthiness

Research have shown that emotions with positive valence such as happiness increases trust while emotions with negative valence, such as anger decreases it. [5] The same brain structures that processes emotion (such as the amygdala), have been indicated to be used for decisions of trustworthiness. [25] Prosocial behaviour, like smiling and mutual gaze are two non-verbal behaviours that have been shown to increase perceived trustworthiness and valence [10][23][13] and that even fake smiling is better than no smiling at all when it comes to trustworthiness [19].

Based upon previous research, 7 different, non-verbal behaviours were constructed and rated in a pre-study (see fig. 6 and fig. 7). The smiling VCs were rated positive in valence but differently in terms of trustworthiness. Two of them showed a tendency of being rated more trustworthy than they were rated in any other value (40% rated them trustworthy). The three VCs with non-verbal behaviours hypothesized to be untrustworthy and negative valence were rated negative in terms of valence but only one had a tendency of being rated as untrustworthy, with 45% of participants rating it as untrustworthy and 7% as very untrustworthy. The last VC with a hypothesized neutral behaviour were rated as neutral in terms of valence and had a tendency of being rated as neutral in terms of trustworthiness (but only by 50%). The VCs were rated clearly in terms of valence, 3 negative, 1 neutral and 3 positive but ratings of trustworthiness were not as clear. This indicates that valence can not be directly used to increase or decrease trustworthiness.

Although valence does not seem to have a direct influence over trustworthiness, trustworthiness seem to have some influence on valence. Looking at the valence ratings before and after the trust game (see fig. 10), the valence ratings changed for the VCs with incongruent personalities. If the VC was perceived as more trustworthy, (receiving more credits in the initial bet) but didn't live up to the expectations, they were rated as less positive in terms of valence. If the VC was perceived as less trustworthy (receiving less credits in the initial bet) but made trustworthy decisions, they were rated as more positive after the trust game. VCs with a congruent personality were unaffected. This could be related to research where attractive counterparts receive more in the initial bet but are penalised if they don't live up to

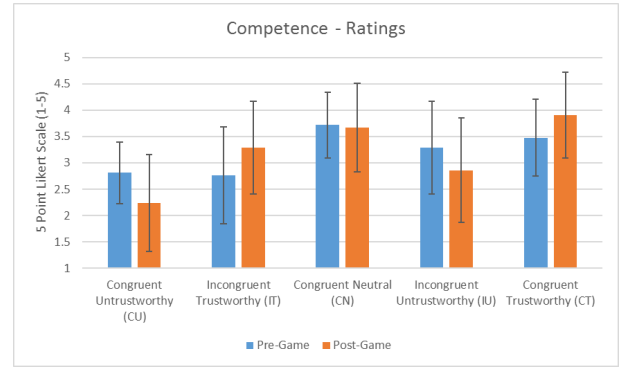


Figure 14: The 5 VCs rated before and after the trust game in terms of competence

the expectations. [24]

The 3 VCs chosen for the main study was rated again prior to the trust game and after the trust game. These were chosen because they had a tendency of being rated as untrustworthy, neutral and trustworthy. Result from the ratings prior to the trust game yielded similar results as the pre-study: no clear differences in trustworthiness but 3 distinct different VCs in terms of valence. This again indicates that valence have no direct influence over trustworthiness.

Non-verbal behaviours like smiling, anger and mutual gaze seems to influence both trustworthiness and valence but that valence itself is no indicator of trustworthiness.

5.2 Trustworthiness in the trust game

Seeing how there was a tendency of VCs being rated as more trustworthy, neutral or untrustworthy in the pre-study (see fig. 7), an interesting factor to investigate is the initial bet in the trust game since it has been shown that the investment in the trust game can be directly linked to perceived trustworthiness. [22] Looking at the initial bets (see fig. 15), the VCs that had a tendency of being rated more trustworthy in the pre-study (IU and CT) received more credits than the VCs with a tendency of being rated as more untrustworthy (CU and IT) which indicates a difference in trustworthiness. Mutual gaze and smiling seems to indicate a more trustworthy VC than averted gaze and nervousness. Interestingly, the VC with a tendency of being rated as more neutral received roughly the same initial amount in the trust game as the VCs with a tendency of being rated as more trustworthy which would indicate that there were no difference in perceived trustworthiness between them.

Judging from the initial bets related to perceived trustworthiness, the hypothesized neutral and hypothesized trustworthy VCs (CN, CT and IU) were perceived similarly but as more trustworthy than the hypothesized untrustworthy VCs (IT and CU).

Interesting to note here is that there was very little previous knowledge into what makes a subject perceived as neutral in the trust game scenario. Previous studies suggest that subjects that give back more or just as much as is given, is considered to be trustworthy, and other subjects are untrustworthy. From initial feedback during the pilot study, a VC giving back between 100 and 120% was not perceived as trustworthy rather, it was perceived as neutral which can be seen in the ratings of trustworthiness (see fig. 9). Previous

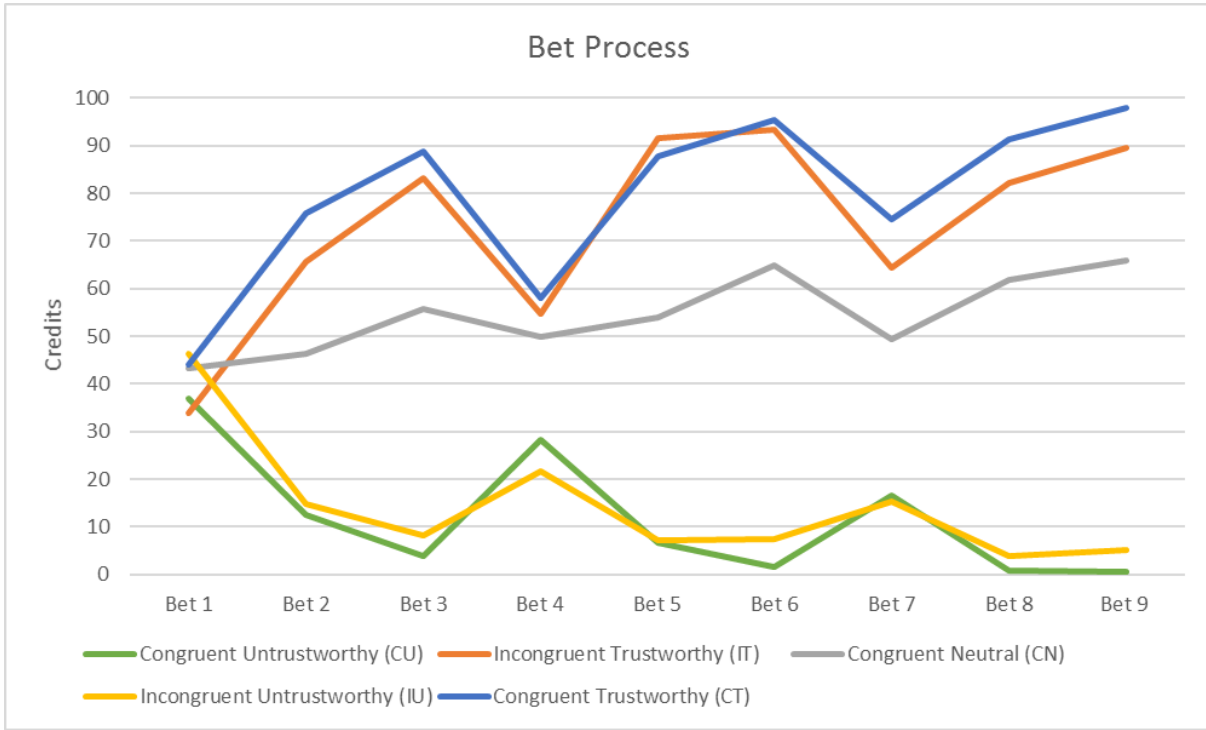


Figure 15: The bet progression (in average) for the 5 VCs

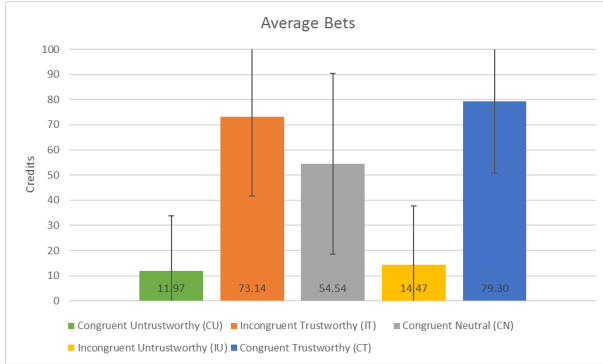


Figure 16: The average bet (investment) given to the VCs in the trust game

research have not introduced neutralness to the trustworthiness ratings and have only asked if counterparts are untrustworthy or trustworthy and to what degree. This could be a reason as to why the results were different to this study.

Non-verbal behaviours like smiling and mutual gaze seem to have an influence on trustworthiness, similar to previous studies [10][23][13][19], but it was not enough to make clearly trustworthy or untrustworthy VCs in this study.

Asking for perceived trustworthiness seems to be a hard task and ratings were mostly ambiguous, where participants rated VCs very differently. Introducing the trust game and measuring trustworthiness by looking at how much participants invested in VCs was a better indicator and showed more clear results.

5.3 Trustworthiness and context

Trustworthiness seems to be hard to judge from a first impression but after adding context, in this case through interaction in a trust game, participants can easier rate VCs in terms of trustworthiness (see fig. 9) and these ratings seem to correlate to perceived altruism (see 11). But are the previous decisions the only factor influencing how much to give in the trust game? Looking at the bet process (see fig. 15) and average bets (see fig. 16), the congruent trustworthy (CT) VC received on average the most credits and the biggest last average investment, followed by the incongruent trustworthy (IT) VC. The congruent untrustworthy (CU) VC received on average the least credits and the lowest last average investment, followed by the incongruent untrustworthy (IU) VC. These results indicate that the non-verbal behaviour of the VCs had an impact after interaction was introduced in the trust game, even in the last and ninth investment. The question whether the VC makes trustworthy or untrustworthy decisions in an interaction does not seem to be the sole basis for trusting it or not, but the VC's non-verbal behaviour still impacts the perception.

5.4 Perception of virtual characters

Can trustworthiness be differently perceived in VCs than how it is perceived in humans? Research have shown that for gaze, they are perceived in the same way [2], but for the overall impression of trustworthiness, this might not be true. Participants in this study was instructed that they were going to play the trust game with actual virtual characters and not humanly-controlled characters (wizard of oz). Much of previous research into trustworthiness with computer generated faces and characters have instructed the participants that they are playing against humanly-controlled characters

[10][20]. This can have an impact on the perception of the counterpart. VCs does not directly profit from monetary gain or prosocial behaviour, neither are they complex human beings controlled by emotion. They are controlled by an algorithm set up by humans and follow this logic completely. This could influence the interaction and decisions made in the trust game.

5.5 Nervousness

Of the 3 non-verbal behaviours used in the main study, the hypothesized untrustworthy non-verbal behaviour included traits of nervousness. The VC had an increased amount of micro and macro eye saccades, little eye-contact, hand gestures around the facial area (right hand scratching neck every now and then) and a confused facial expression (see fig. 1). From the results (see fig. 12), the VCs with this non-verbal behaviour (CU and IT) were perceived as nervous. They were also perceived as negative in terms of valence (see csfig. 10). This indicates that these non-verbal behaviours can be used to create a VC that is perceived as nervous and that there is a relationship between nervousness and negative valence and also between confidence and positive valence.

5.6 Critical discussion

To execute this thesis project, knowledge of several complicated areas were encompassed and incorporated in the design. Without being an expert in any of these domains there was a big a challenge in executing the work well. One of these complicated areas was gaze related behaviour. The VCs that participants played the trust game with used an algorithm to decide whether to look the participant in the eye or not. This was decided with a percentage chance. The hypothesized trustworthy characters had a 80 percentage chance of mutual gaze and the hypothesized untrustworthy and neutral had a 30 percentage chance. The duration for mutual gaze lasted between 2 and 6 second and then a new roll of the die would decide if the VC would continue to maintain mutual gaze or look away. The micro and macro eye saccades and blinking worked in a similar way. Looking at it from a strict experimental point of view, as many of these loose variables should be controlled for and it would have been safer to have static durations for everything to get results with even more credibility. The decisions made to keep it this way was based on that it was more important for the VCs to have unique behaviours that could not be learnt from a longer interaction. It was more important that the participants perceived the virtual characters as human-like and intelligent. Also, similar studies have used percentages to manipulate mutual gaze and found that participants can tell the difference between mutual gaze and averted gaze using this method. [13]

In the pre-study, 4 different types of stimuli were being investigated but these were not divided into 4 individual blocks. All stimuli were counterbalanced together which might have had an impact on the results, although, no significant change in ratings of trustworthiness could be found between the pre-study and main study (see fig. 8).

When participants familiarized themselves with the trust game in the main study, they were playing with a random character from the pool of 5 virtual characters (excluding the one they were going to start playing with when the test began for real). This could have had an impact on the initial

impression and ratings of the VCs because every participant had interacted with one of them prior to the ratings.

Moreover, inviting more participants is always something that would increase credibility. The participants in the pre-study and main study were all students at the Royal Institute of Technology but not everyone was Swedish. Previous research [9] suggests a difference in interpreting emotions and trust differently throughout different cultures and this could have had an impact on the results. Nothing could be found by dividing participants in different culture groups though.

6. CONCLUSION

Valence seem to have no direct influence over perceived trustworthiness based on the results in this study. Non-verbal behaviours that influence positive valence, such as smiling and mutual gaze seem to have an influence over trustworthiness. Interestingly the neutral virtual character with little mutual gaze and no facial expression received just as much credits in the initial investment in the trust game as the virtual characters that were smiling and maintaining a high percentage of mutual gaze which indicates that they were perceived the same in terms of trustworthiness. Constructing a non-verbal behaviour based on valence and mutual gaze alone does not seem to be enough to create trustworthy virtual characters. Moreover, the non-verbal behaviour of a virtual character will still influence the ninth and last bet in the trust game, indicating that it is not only the previous decisions of trust of the virtual character that will influence how trustworthy it is perceived.

6.1 Future Research

The difference of perception between humans and virtual characters are of importance in a world where virtual characters are becoming more and more common interaction partners. Studies have shown that there is no difference in perception of gaze for humans and VCs [2] but there are few studies investigating the difference of other factors such as valence and facial expressions. One might think that there is a difference, that complex human beings are perceived differently from static, logic-following VCs - but this needs to be investigated further.

Another interesting future investigation is to continue in this study's steps and explore how to create a really trustworthy VC. A recipe for a non-verbal behaviour that could be applied to any, or most, appearances would be of incredible use when designing for training, education, healthcare or even in other areas where user experience is of importance.

Little research could also be found on the impact body motion have on perception of trustworthiness. It was hard to know how this would influence and to what degree. Investigating the relationship between body motion and trustworthiness in a scenario involving VCs would be useful.

7. ACKNOWLEDGEMENTS

I would like to thank Robin Palmberg, a fellow master thesis student, who I worked closely together with during the last months of the thesis project and who have given valuable feedback on this report. I would also like to thank my supervisor, Christopher Peters, who have helped guide this thesis from square one and have always been there to give advice and feedback.

8. REFERENCES

- [1] A. Abdul-Rahman and S. Hailes. Supporting trust in virtual communities. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, volume vol.1, page 9. IEEE Comput. Soc, 2000.
- [2] S. Andrist, T. Pejisa, B. Mutlu, and M. Gleicher. Designing effective gaze mechanisms for virtual agents. *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, pages 705–714, 2012.
- [3] J. Berg, J. Dickhaut, and K. McCabe. Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 10(1):122–142, jul 1995.
- [4] J. Cassell and K. R. Thorisson. The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence*, 13(4-5):519–538, 1999.
- [5] J. R. Dunn and M. E. Schweitzer. Feeling and Believing: The Influence of Emotion on Trust.
- [6] R. S. Gowda. Role of Software Agents in E-Commerce. *International Journal Of Computational Engineering Research*, 3(3):246–251, 2013.
- [7] E. Hehman, J. K. Flake, and J. B. Freeman. Static and Dynamic Facial Cues Differentially Affect the Consistency of Social Evaluations. *Personality and Social Psychology Bulletin*, 41(8):1–12, aug 2015.
- [8] D. Keltner, A. Kogan, P. K. Piff, and S. R. Saturn. The sociocultural appraisals, values, and emotions (SAVE) framework of prosociality: core processes from gene to meme. *Annual review of psychology*, 65:425–60, 2014.
- [9] A. Kleinsmith, P. R. De Silva, and N. Bianchi-Berthouze. Cross-cultural differences in recognizing affect from body posture. *Interacting with Computers*, 18(6):1371–1389, 2006.
- [10] E. G. Krumhuber, A. S. R. Manstead, D. Cosker, D. Marshall, P. L. Rosin, and A. Kappas. Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion*, 7(4):730–5, 2007.
- [11] J. C. Lester, S. G. Towns, C. B. Callaway, J. L. Voerman, and P. J. FitzGerald. Embodied conversational agents. pages 123–154. MIT Press, 2000.
- [12] D. H. Mcknight and N. L. Chervany. The meanings of trust. *Measurement*, 55455(612):86, 1996.
- [13] A. Normoyle, J. B. Badler, T. Fan, N. I. Badler, V. J. Cassol, and S. R. Musse. Evaluating Perceived Trust From Procedurally Animated Gaze Evaluating Perceived Trust From Procedurally Animated Gaze Evaluating perceived trust from procedurally animated gaze. 2013.
- [14] N. N. Oosterhof and A. Todorov. The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(32):11087–92, aug 2008.
- [15] N. N. Oosterhof and A. Todorov. Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion (Washington, D.C.)*, 9(1):128–133, 2009.
- [16] J. Rickel and W. Johnson. Task-oriented collaboration with embodied agents in virtual worlds. *Embodied conversational agents*, pages 1–29, 2000.
- [17] D. Rigas and N. Gazepidis. A further investigation of facial expressions and body gestures as metaphors in e-commerce. In *Proceedings of the 7th Conference on 7th WSEAS International Conference on Applied Informatics and Communications*, pages 148–153, 2007.
- [18] D. M. Rousseau, S. B. Sitkin, R. S. Burt, and C. Camerer. Not so different after all: a cross-discipline view of trust. *Academy of Management Review*, 23(3):393–404, jul 1998.
- [19] J. P. Scharlemann, C. C. Eckel, A. Kacelnik, and R. K. Wilson. The value of a smile: Game theory with a human face. *Journal of Economic Psychology*, 22(5):617–640, 2001.
- [20] M. Stirrat and D. I. Perrett. Valid facial cues to cooperation and trust: male facial width and trustworthiness. *Psychological science*, 21(3):349–54, mar 2010.
- [21] A. Tartaro and J. Cassell. Authorable virtual peers for autism spectrum disorders. In *Proceedings of the Combined workshop on LanguageEnabled Educational Technology and Development and Evaluation for Robust Spoken Dialogue Systems at the 17th European Conference on Artificial Intelligence*, volume 174, pages 169–174, 2006.
- [22] M. van 't Wout and A. G. Sanfey. Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition*, 108(3):796–803, sep 2008.
- [23] M. L. Willis, R. Palermo, and D. Burke. Social Judgments are Influenced By Both Facial Expression and Direction of Eye Gaze. *Social Cognition*, 29(4):415–429, 2011.
- [24] R. K. Wilson and C. C. Eckel. Judging a Book by its Cover: Beauty and Expectations in the Trust Game. *Political Research Quarterly*, 59(2):189–202, jun 2006.
- [25] J. S. Winston, B. A. Strange, J. O 'doherty, and R. J. Dolan. Automatic and intentional brain responses during evaluation of trustworthiness of faces. 2002.

