DEGREE PROJECT IN MEDIA TECHNOLOGY,
SECOND CYCLE, 30 CREDITS
STOCKHOLM, SWEDEN 2016

The Effect of Facial Expressions Valence on the Perception of the Body Motions of Virtual Groups

ROBIN PALMBERG

KTH ROYAL INSTITUTE OF TECHNOLOGY
SCHOOL OF COMPUTER SCIENCE AND COMMUNICATION

The Effect of Facial Expressions Valence on the Perception of the Body Motions of Virtual Groups

Effekten av Känslomässiga Ansiktsuttryck på Uppfattningen av Kroppsrörelser hos Virtuella Grupper

Robin Palmberg
robinpa@kth.se

Master's Thesis in Media Technology (30 credits)
at the School of Computer Science and Communication
Royal Institute of Technology year 2016

Supervisor was Christopher Peters

Supervisor was Roberto Bresin

2016-10-03

ABSTRACT

To understand what modality affects the perception of emotion is interesting since it would help us get closer to understanding and dissecting emotions. Earlier research has shown that body motions can help disambiguate ambiguous (determining something vaguely presented) facial expressions but there is no study that examines the possibility of facial expressions to affect the perception of body motion. That is why this study aims to examine if facial expressions can affect the perception of full body emotions. This is done with a perceptual experiment in which 22 subjects were exposed to stimuli consisting of scenes with virtual characters expressing emotions (see figure 1). It was concluded that the facial expression does affect the perception of joined valence within the group of characters and it is thereby proven that not only body motions can help disambiguating ambiguous facial expressions but that facial expressions can help disambiguating ambiguous body motions and alter the perception of distinct body motions perceived as either positive or negative without face and hands showing. It is also concluded that perceived trustworthiness is affected by the valence of the facial expression, which supports recent studies conducted which aimed to find out what makes for a trustworthy and dominant face with the help of valence as a factor. The perceived relationship within the group is also examined as well as the dominance of the group as a whole but neither gave results clear enough to make any conclusions except that positive valence affects the perceived relationship within the group as closer. The study is conducted using virtual agents but is meant to help better understand people in everyday situations.

SAMMANFATTNING

Att förstå vilka modaliteter som påverkar uppfattningen av känslor är intressant eftersom det skulle kunna hjälpa oss att komma närmare att förstå och kunna analysera känslor samt deras beståndsdelar. Tidigare forskning har visat att kroppsrörelser kan hjälpa till att förtydliga tvetydiga ansiktsuttryck men det finns ingen studie som undersöker möjligheten att ansiktsuttryck skulle kunna påverka uppfattningen av kroppsrörelser. Därför har denna studie som mål att undersöka om ansiktsuttryck kan påverka uppfattningen av känslor uttryckta med helkroppsrörelser. Detta undersöks med ett perceptuellt experiment där 22 försökspersoner exponerades för stimuli som bestod av scener med virtuella karaktärer som uttryckte känslor (se figur 1). Det konstaterades att ansiktsuttryck påverkar uppfattningen av den gemensamma valensen i gruppen av virtuella karaktärer och det är därmed bevisat att inte bara kroppsrörelser kan hjälpa till med att förtydliga tvetydiga ansiktsuttryck utan att ansiktsuttryck kan hjälpa till med att förtydliga tvetydiga kroppsrörelser och förändra uppfattningen av distinkta kroppsrörelser som uppfattats som antingen positiva eller negativa utan att ansiktet och händerna visades. Det konstaterades även att den uppfattade trovärdigheten påverkas av ansiktsuttryckets valens, vilket stödjer tidigare studier som hade som mål att undersöka vad som uppfattas som ett trovärdigt och dominant ansikte med hjälp av valens som en faktor. Den uppfattade relationen inom gruppen av karaktärer är också undersökt samt dominansen av gruppen som helhet över försökspersonen men ingen av dessa gav resultat som var tydliga nog för att kunna dra några slutsatser förutom att positiv valens (samling av positiva känslor) påverkar den uppfattade relationen i gruppen som närmare. Studien utfördes med användandet av virtuella agenter men är tänkt att kunna hjälpa oss att bättre förstå människor i vardagssituationer.

1. INTRODUCTION 1
2. THEORY AND RELATED RESEARCH 2
2.1 Research question and hypothesis 2
3. METHOD 2
3.1 Technology 2
3.2 Prestudy 3
3.2.1 Stimuli and questions 3
3.3 Main Study 4
3.3.1 Stimuli and questions 5
3.3.2 Connection to other study 5
4. RESULTS 5
4.1 Prestudy 5
4.1.1 Virtual character appearance 5
4.1.2 Facial static expressions and body motions 5
4.2 Main Study 6
4.2.1 Valence 6
4.2.2 Trustworthiness 7
4.2.3 Relationship 7
4.2.4 Dominance 7
5. DISCUSSION 7
5.1 Valence 8
5.2 Other factors 8
5.2.1 Trustworthiness 8
5.2.2 Relationship 8
5.2.3 Dominance 8
5.3 In relation to other studies 8
5.4 Limitations and considerations 8
6. APPLICATIONS 9
7. CONCLUSION 9
7.1 Future work and improvements 9
8. ACKNOWLEDGEMENTS 9
9. REFERENCES 10

Figure 1 - A group of virtual characters interacting, this study concerns the way in which facial and body motions contribute to our perception of emotions

CCS Concepts

• Computing methodologies~Physical simulation
• Computing methodologies~Perception

Keywords

Body motion; Facial animation; Perception; Virtual Characters

1. INTRODUCTION

What affects the perception of an emotion while looking at groups of people? This is an important question since the knowledge about this could help people in many situations. There are a lot of things that combined creates the emotions that we perceive, but what modalities are important when trying to perceive an emotion? In society today it makes all the difference to be able to recognize the emotion a person has to be able to analyze a situation quickly and adapt to it. With more knowledge about how the emotions are perceived we can learn to be better at understanding each other.

One way of investigating how emotions are perceived is by using virtual agents. Virtual agents are becoming a growing part of our everyday life. They are encountered in (amongst other) video games, movies, websites and receptions and can be in the form of a virtual character displayed on a screen or as a social robot that can be interacted with by human beings. They can also help with human computer interaction for those users who does not have the ability to use conventional ways of interaction. A study conducted by Lovato and Piper [12] examines how children who can neither read nor write interact with computers by using vocal input and receiving audible feedback.

Virtual characters are often created in such a way as to mimic real life human beings and that can help in science since you are able to perform and conduct experiments in new manners using these virtual characters as "mannequins" instead of having to use real actors and the likes of it. These experiments can be based in psychology, computer science and other research areas since they are versatile and cheap to use. One example is the study conducted by Beck, Cañamero, Hiolle et al. [2] where social robots are used to display body language to find out how humanoid robots body language is perceived by children.

Games for teaching emotions have also been made for children [10]. This is yet another example on how virtual characters can be used in other ways than the conventional ones, to convey emotions and let children learn from them. As an example of application it has been done with children diagnosed with autism with good results [5].

As mentioned earlier an emotion consists of several modalities which together creates the overall perceived emotion. But is any of the modalities salient (more obvious than the others) or is the combination the only thing that gives us the full information? The study conducted by Ennis, Hoyet, Egges et al. [8] concludes that the modalities combined gives a better effect but that they can be determined on their own if needed.

Some emotions can be ambiguous but intense, which means that you can see that a person is expressing something very much but you cannot say what it is. The aim of this study is to conclude if and how facial expressions affect the overall perception of an emotion with both a body motion that is ambiguous and body motions which are annotated as either positive or negative. This is done with the use of virtual agents in a game engine, which are presented as video clips in a study software that collects data as well. The results show that facial expressions alter the perception of valence when looking at a group expressing different combinations of body motions and facial static expression which are either congruent (face and body expresses the same emotion) or incongruent (face and body expresses emotions that are not related, for example: sad body with a happy face).

This document consists of a literature study presented in "2. Theory and related research". Then the methods chosen are explained in "3. Method" after which the results are presented in "4. Results". The discussion in "5. Discussion" leads in to ways of implementation in "6. Applications". Lastly a conclusion is made in "7. Conclusion" and after that "8. Acknowledgements" were gratitude for all the help is presented.

2. THEORY AND RELATED RESEARCH

The significance and saliency of modalities are significant and in a recent study conducted by Aviezer, Trope and Todorov [1] they concluded that body language could help disambiguate ambiguous facial expressions. The study was done using photographs of distinct body motions and combined them in a photo editing tool with different ambiguous facial expression which were congruent in the way that they had the same physical aspects as the body (they did not use different genders for face and body as that would be incongruent and could disturb the study). This study of course raises the question about how facial expressions could affect the perception of ambiguous body motions since it is quite the opposite of the conclusion made in that study. Also the study never touched the topic of congruent and incongruent facial-body relations rather than just assuming that a full body emotion would be analyzed by the brain at the same time.

A study conducted by Meeren, van Heijnsbergen and de Gelder [14] handle that question concerning how we perceive congruent and incongruent facial-body relations. The study tells us that before the brain has had time to analyze the emotional state of an individual it already categorizes the facial-body relation as either congruent or incongruent. This raises the question of what parts of the individual that the brain focuses on in that case. It also raises a question about if there is any modality of the emotion that is salient and more important than the others to be able to analyze an emotion. According to Dahl and Friberg [7] different parts of the body is analyzed with just slight influence depending on what part of the body that is being displayed.

With these two studies in mind a research question starts to take shape. If there is a modality that is salient, which would it be? The study [1] did claim that the body motion could help disambiguate ambiguous facial expressions. Would it be possible to turn the tables and see if the facial expressions could help disambiguate ambiguous body motions? Oosterhof and Todorov [15] show that people are able to perceive valence and dominance through facial expressions and facial cues. With this knowledge it seems quite possible that the facial expression could affect the perception of an ambiguous body motion, especially since [14] have proven that the brain analyses the modalities by themselves after 120 milliseconds if the facial-body relation is incongruent.

A factor that might have disturbed the results of the study by [1] is that they kept the backgrounds for each of the body motions. As Kret and de Gelder [11] state we already know that the perception of facial expressions are affected by the surroundings [4] and that is why they conducted a study to find out if the perception of the body motion was affected by the background as well and they found out that it at least had an effect on the time it took to analyze the emotion of the person of interest.

Previous studies have therefore shown that the background and the situation of the setting that the virtual agents are in do affect the perception of emotions expressed by the characters. In other studies, evaluating the perceived emotion of virtual characters the background was used in such a way as to create a neutral setting in which they could test how characters in the background affect the perception of the characters in focus. This study was conducted by Carretero, Qureshi and Peters [3] and it looks a lot like the study [11]. However instead of photographs androgynous virtual dolls (genderless and faceless dolls) were used and they were used in moving pictures instead of just static images. This study again finds that background situations affect the perception of emotions of the virtual characters.

As it seems there have not been much work made on isolated people expressing full body motions. This is an interesting topic since it would give information that would be transferrable from situation to situation.

Another interesting factor when working with the perception of emotions is how the emotions are perceived when put in a group context. The study conducted [3] focused on a group standing in the foreground of a scene and a study conducted by Ennis, McDonnel and O'Sullivan [9] examined what makes a group interaction feel realistic. In the experiment they instead of using a background used a simple, plain white background, isolating the group from everything else that could disturb the results.

2.1 Research question and hypothesis

With these studies in mind we can find that there is a gap where one could study how a groups joined emotion is perceived in an isolated environment and by doing so also find out if there is any modality that has a certain significance or possibility to affect the overall impression that is perceived by the user. This makes for a study which could be conducted to find out how different parts of the body affects the perception of a groups joined emotion.

To be able to conduct this experiment annotated facial expression, body motions and appearances is needed. These could then be used to conduct a study which answers the question:

Does a facial static expression affect the way body motion are perceived in a group scenario?

Because of the findings in [14] there is reason to believe that facial expression could be more important than body motions when determining expressed emotions. The hypothesis of this study is that the facial expression will affect the overall perception of emotions in a group where all the characters are expressing the same emotion.

3. METHOD

A prestudy and a main study were conducted, each with 22 participants from which the data of 21 could be used since one participant's set of data had to be discarded in both of the studies.

3.1 Technology

The same technology and method was used in both cases, the only factor that changed was what stimuli that was used. The stimuli consisted of scenes created in a game engine called *Unity3D*[1] using assets bought through the *Unity Asset store*[2]. The game engine is free to use but most of the assets had to be bought and were payed for by the institute. For the virtual character *MCS Female*[3] by *Morph3D* was used in combination with their clothes packs *MCS Female: Jersey Girl*[4] and *MCS Female: Sci-Fi Bundle*

[1] https://unity3d.com/

[2] https://www.assetstore.unity3d.com/

[3] https://www.assetstore.unity3d.com/en/#!/content/45807

[4] https://www.assetstore.unity3d.com/en/#!/content/51541

01[5]. For body animations the *Social Motion Pack Take 1*[6] by *PolygonCraft* was used in combination with the animations from *Taichi Character Pack*[7] by *Game Asset Studio* and *Realistic Eye Movements*[8] by *Tore Knabe*. Each scene was built as a standalone, full screen executable file which was later recorded with *ZD Soft Screen Recorder*[9] with a resolution of 1920x1080 pixels and a frame rate of 30 frames per second.

The recorded scenes were put in the software *Tobii Studio Pro*[10] that displayed the stimuli in a certain way of randomization that is called "latin square"[11]. For the experiment each video was displayed approximately 10 seconds and followed by a set of questions. The data was exported and analyzed using functions in *Microsoft Excel*. The screen used for the study was a 24 inch LED-screen with 1920x1080 pixels resolution and the participants were sitting approximately 60 cm from the screen. All interaction with the software by the participant was made with a computer mouse.

In both the prestudy and the main study the body animations were created using the animator controller built into Unity3D and the facial static expressions were created using the MCS Females blend shapes. In the prestudy all of the facial expressions were shown with the same body motion (an idle motion that were not used in the main study) and all of the body motions had blurred faces and hands. These two components were later put together to create the stimuli for the main study by putting the animator controller connected to the MCS Female model and using the blend shapes that were built into the MCS Female model.

3.2 Prestudy

The prestudy was conducted to annotate the perceived valence of several virtual characters, animations and facial expressions. This information was used to select which of these that would be suitable for the main study. The prestudy was conducted together with another master thesis student who was focusing on the trustworthiness of virtual characters. The prestudy took about 30 minutes per participant and there were almost always two participants doing the study at the same time in the same room sitting back to back with one master thesis student as a supervisor each.

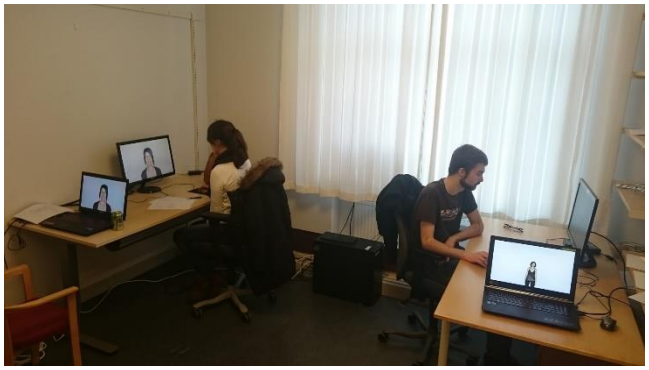

Figure 2 - Two participants taking part in the prestudy

As for the introduction of the study each participant got to read a text consisting of the definitions of valence and trustworthiness as well as a scenario for the experiment which read: *Imagine you are playing a game with a virtual character. The first step of the game is to choose a virtual character to cooperate with. You will be rating the virtual character based on these two factors: Trustworthiness – How much do you think you can trust the virtual character? How trustworthy is the virtual character? Valence – Is the emotion expressed by the virtual character in the video clip a positive, a neutral or a negative one?*

Each of the participants got to run the test two times to get more samples per participant. In the beginning of each test the new participant was displayed with 3 dummy stimuli which they could answer anything on and which they could ask questions about. As soon as the test started they were not allowed to ask any more questions. As a reward each participant was given cookies and buns.

3.2.1 Stimuli and questions

There were 23 stimuli that were presented: 7 different appearances (see figure 3 for example), 7 "behaviors" (see figure 4 for example) (combinations of facial animations and idle body motions) (not used in this study), 5 static facial expressions in combination with an idle animation (see figure 5 for example) and 4 body motions (consisting of strings of 2-3 animations each) where the face and hands are blurred as to not affect the valence of the body motions (see figure 6 for example). The purpose when blurring the face and hands was to eliminate any affect the face could have on the perception of the expressed emotion. Gestures made by the hands may be interpreted in a way that affects the perception of the full body motion, for example a closed fist might display anger or frustration even though the rest of the body is displaying a positive emotion.

The static facial expressions were created by using different factors defined by the company who had designed the model. They had their own annotations for the factors and the ones used for this study was squinting eyes, smiling mouth and the feelings happiness, sadness, shock and surprise.

Each questionnaire consisted of two questions, valence (on the scale: *very negative, negative, neutral, positive, very positive*) and trustworthiness (on the scale: *very untrustworthy, untrustworthy, neutral, trustworthy, very trustworthy*).

[5] https://www.assetstore.unity3d.com/en/#!/content/45811

[6] https://www.assetstore.unity3d.com/en/#!/content/26099

[7] https://www.assetstore.unity3d.com/en/#!/content/15667

[8] https://www.assetstore.unity3d.com/en/#!/content/29168

[9] https://www.zdsoft.com/

[10] http://www.tobiipro.com/product-listing/tobii-pro-studio/

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

Figure 3 - Example of stimuli, this is one of the appearances

Figure 4 - Example of stimuli, this is one of the behaviors

Figure 5 - Example of stimuli, this is a positive facial static expression

Figure 6 - Example of stimuli, this is body motion 3

3.3 Main Study

The main study was conducted in the same way as the prestudy in the sense that Tobii Studio Pro was used to display stimuli and collect data from the participants but instead of the old stimuli there were now 12 new stimuli. These were based on the results of the prestudy as the prestudy concluded that there were three distinct facial expressions that were displaying either negative, neutral or positive valence as well as three body motions. The body motions were perceived differently, there was two distinct and one ambiguous body motion. The distinct body motions displayed either positive or negative valence while the perception of the ambiguous body motion was really hard do decipher, therefore it is treated as an ambiguous body motion. A fourth factor was added by also using blurred faces and hands.

As for the introduction of the study each participant got to read a text consisting of the definitions of valence and trustworthiness as well as a scenario for the experiment which read: *You are seeing a group of characters interacting with each other in a game. You are not sure if you should approach them or not. To make your decision you will rate them according to these four factors:*

Valence (very negative-negative-neutral-positive-very positive) - Is the emotion expressed by the virtual character in the video clip a positive, a neutral or a negative one? A positive emotion could be joy or happiness while a negative emotion could be sadness or anger.

Trustworthiness (very untrustworthy-untrustworthy-neutral-trustworthy-very trustworthy) - How much do you think you can trust the virtual character? How trustworthy is the virtual character?

Relationship (very distant-distant-neutral-close-very close) - What kind of relationship do these characters have to eachother?

Dominance (very inferior-inferior-neutral-dominant-very dominant) - How dominant are the characters in the group as a whole?

For the main study the focus was on the perceived valence, trustworthiness, relationship and dominance within the group. The valence and trustworthiness had the same scales as in the prestudy but the relationship had the scale *very distant, distant, neutral, close, very close* while the dominance had the scale *very inferior, inferior, neutral, dominant, very dominant.* All of the questions were showed in the same order after each of the video clip.

Each participant got to run the test 4 times each as to be able to collect more samples per participant and get better and stronger results. The test was divided in blocks which consisted of each stimulus presented once. In the beginning of each experiment once the participant had joined the session the new participant was

displayed with 3 examples which they could answer anything on and which they could ask questions about. As soon as the real test started they were not allowed to ask any more questions. The experiment had a duration of 30 minutes. As a reward, the participants were given candy.

3.3.1 Stimuli and questions

The stimuli were created using the valence data from the prestudy. The 3 most neutral characters/appearances were chosen, put in a group scenario (see figure 7) and used for expressing a positive, a negative and an ambiguous body motion. Each of these body motions were ran four times, once with a positive, a negative, a neutral and a blurred static facial expression (for the blurred face the hands were also blurred as to not affect the perception of the body motion) (see table 1).

	Blurred face	Negative face	Neutral face	Positive face
Negative body motion	Scene 1	Scene 2	Scene 3	Scene 4
Neutral body motion	Scene 5	Scene 6	Scene 7	Scene 8
Positive body motion	Scene 9	Scene 10	Scene 11	Scene 12

Table 1 - The combination of facial static expressions and body motions and which scenes they create when combined

Figure 7 - Example of what a scene looks like

The body motions consisted of sets of segments of full body animations which were combined to create full body emotions. Each character started at a separate segment of the animations in the same body motion as the other characters as to not make the characters show the exact same animations at the same time. When the segments of animations were put together they had a total duration of approximately 10 seconds.

The body motions were put in a group context (see figure 7) and the scenes were combined as table 1 shows. These 12 scenes were put in a semi-random order using "latin square" which is a way of randomizing a number of stimuli that is going to be shown several times so that no randomization is the same as the others. This was done as to avoid patterns of stimuli since that could be a factor that would affect the perception the stimuli.

3.3.2 Connection to other study

As a way of getting more participants this main study was made in conjunction to another main study conducted by master student Niklas Blomqvist. First each participant was part of this study and they were then transferred to another room to take part in Niklas' experiment after which they filled in a questionnaire about themselves which were used by both studies. Together, these studies had a duration of 45-60 minutes.

4. RESULTS

4.1 Prestudy

The prestudy gave results which showed what characters, animations and facial expressions that should be used in the main study. 22 participants (17 men and 5 women) did the prestudy all of which were students, most of them with a focus on computer science at either KTH or Stockholms University. However, one of the participant's results had to be crossed out because of misunderstandings. 20 of the remaining 21 participants were in the age range of 20-30 years old and 16 of them were men and 5 were women.

4.1.1 Virtual character appearance

Firstly, the characters were to be chosen and they should be as neutral as possible as to not affect the perception of the body motion and facial static expressions.

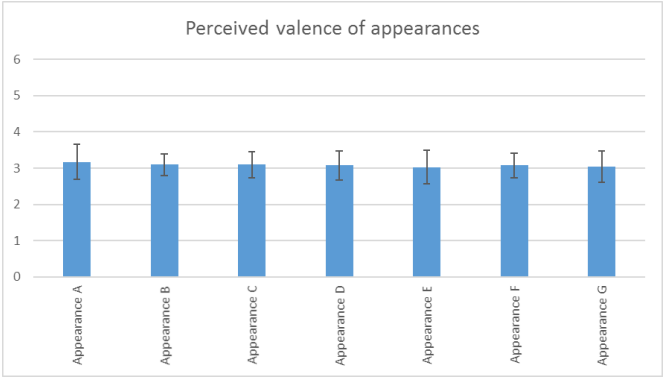

Figure 8 - The perceived valence of the virtual character's appearances

The character appearances B, C and F were chosen for the main study since they were the characters with the most neutral valence and the least standard deviation (see figure 8). The scale of the diagram uses the likert scale that was used in the test. The numbers correspond to the selections 1=very negative, 2=negative, 3=neutral, 4=positive, 5=very positive, this is used in all of the figures relating to valence. The fact that the standard deviation is small means that the participant's responses were quite consistent and that is exactly what is needed for the main study since a selection of precisely annotated stimuli helps making this study credible.

4.1.2 Facial static expressions and body motions

In the same way as the virtual character appearances were chosen a set of 3 facial static expressions (one positive, one neutral and one negative) as well as three body motions (one positive, one neutral and one negative) were to be chosen. The names of the facial static expressions were annotated by this papers author to get a good overview of what they were supposed to show, that must not correlate to the perceived emotion however.

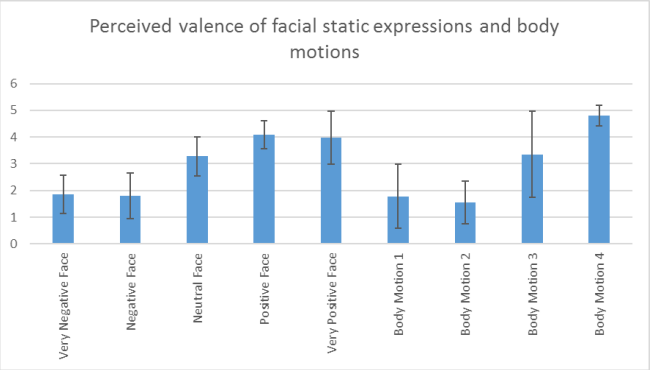

Figure 9 - The perceived valence of the facial static expressions and body motions

It can be seen that the neutral, positive and very negative face should be the most suitable facial expressions since they are the most distinct values with the least standard deviation (see figure 9). An ANOVA test gives the result that the chance of a similarity between these three can be rejected with a 100% cumulative probability.

As for the body motions the body motion 2, 3 and 4 will be chosen since body motion 2 is a distinct negative motion, body motion 3 is an ambiguous motion (as seen by the standard deviation) and body motion 4 is a distinct positive motion (see figure 9). In the same way as with the facial animations an ANOVA test was performed and it gave the result that a similarity between these three body motions could be rejected with a 100% cumulative probablilty.

4.2 Main Study

22 participants (17 men and 5 women) did the main study and each of them did it 4 times which gave a total of 88 responses for each question which is expected to produce a good representation of the perception of the different factors of the test. One participant's responses had to be deleted as it was brought to attention that the participant was diagnosed with Asperger, which is a factor that could affect the perception of displayed stimuli. Out of the remaining 21 participants 19 were in the age range of 20-30 years old and 16 of them were men and 5 were women. 15 of the participants in the main study had also been part of the prestudy.

The participants were presented with a scenario (as explained in the method) which said that they had to determine the valence, trustworthiness, relationship and dominance of this group if they were to meet them in an online game.

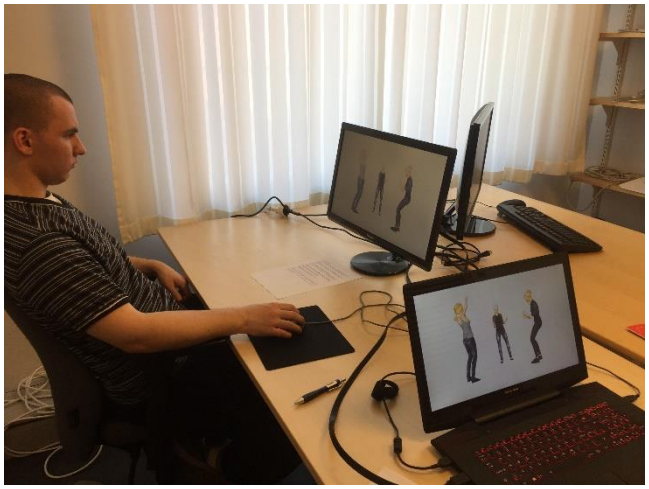

Figure 10 - One of the participants watching one of the video clips in the experiment

4.2.1 Valence

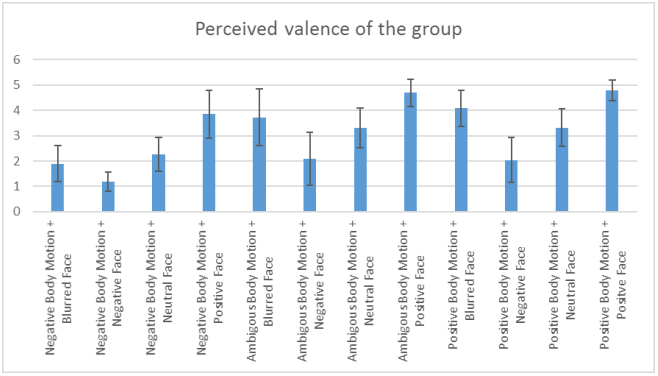

Figure 11 - The perceived valence of the body motions in combination with the facial static expressions

In figure 11 it clearly shows that the perceived valence does change depending on the facial expression. The numbers correspond to the selections 1=very negative, 2=negative, 3=neutral, 4=positive, 5=very positive. The same body motion with the three different faces gives us completely different valences which correlate to the valence of the facial expression. An ANOVA test shows that similarity between the three facial animations on each of the body motions could be rejected with a 100% cumulative probability. However, the blurred face does not give an as clear picture of perceived valence, but it corresponds quite well to the results from the prestudy (see figure 9).

4.2.2 Trustworthiness

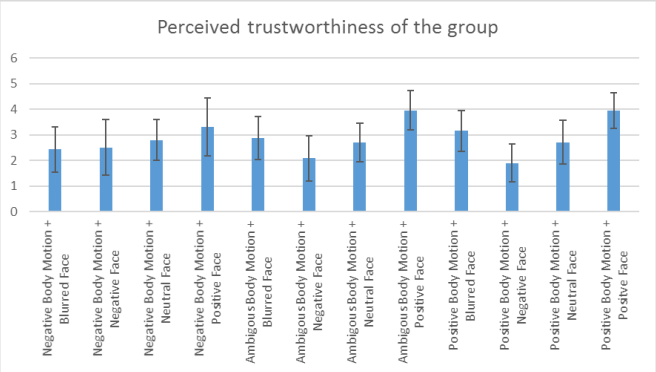

Figure 12 - The perceived trustworthiness of the body motion in combination with the facial static expressions

Just like with the valence the trustworthiness changed depending on the facial animation giving a more trustworthy impression the more positive the facial expression was. The numbers correspond to the selections 1=very untrustworthy, 2=untrustworthy, 3=neutral, 4=trustworthy, 5=very trustworthy. The difference is not as big in figure 12 between the bars as it is in figure 11 but an ANOVA test shows that there is 100% cumulative probability that any similarities between the facial animations of each body motion could be rejected.

4.2.3 Relationship

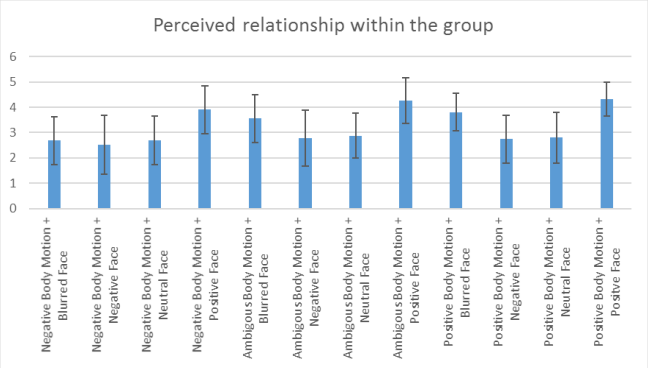

Figure 13 - The perceived relationship between the characters in the group

For relationship the trend is less clear than figure 11 and 12, each answer has quite large standard deviations and it is not possible to make a distinction between the negative and neutral facial expression for any of the body motions (see figure 13) by just looking at it. An ANOVA test shows that a similarity between the facial animations of each body motion could be rejected with 100% cumulative probability. The numbers correspond to the selections 1=very distant, 2=distant, 3=neutral, 4=close, 5=very close.

4.2.4 Dominance

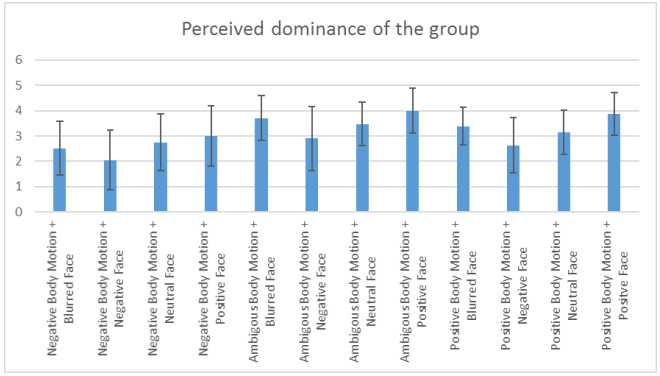

Figure 14 - The perceived dominance of the group as a whole

Just as in the case with the relationship (figure 13), the dominance is quite hard do see a clear trend within (see figure 14). Although an ANOVA test shows that a similarity between the three facial expressions of each body motion could be rejected with 100% cumulative probability. The numbers correspond to the selections 1=very inferior, 2=inferior, 3=neutral, 4=dominant, 5=very dominant.

5. DISCUSSION

The main study gave interesting results, mainly focusing on the valence. The other factors gave results but neither were as clear as the valence and therefore the main focus will be on the valence.

It is clear that the face actually affects the perception of the valence of a group expressing emotions through body motion. Not only do they help disambiguating ambiguous body motion but they also alter the perception of distinct body motions that have been annotated as either positive or negative. A positive body motion with a negative facial expression calls for a perceived negative valence, however not as negative as a negative body motion with a negative facial expression.

The aim of the pre study was to annotate and validate facial expressions, body motions and appearances so that these could be mixed in a way to see if the facial expression alters the perception of the body motion much like how the study by [1] was made. The results were clear and distinct, it was not hard to see which expressions, motions and appearances that should be chosen to be able to study this.

The main study gave clear results which can lead us to believe that the facial expression does change the perception of the body motion no matter the valence of the body motion as long as the expressions are distinct. However, we still do not know if the facial expression is salient or just a part of the full body perception of emotions. The interesting fact that this is done in a group environment also leads us to believe that the joined expression is perceived with a focus on the facial expressions. It would be interesting to compare this to a study which used the same set up but with one character instead of three.

In this set up the stimuli were quite small in relation to the participant's retina. If this affected the results in any way is hard to tell but to be sure the same study could be conducted with the use of a projector or larger screen to get a more realistic size of the characters and also a better stimulus to retina size rate. However, this is discussed in the limitations and considerations section (5.4) but should not make a difference in the scientifically impact as this is the same case as watching a group from a distance.

The four factors examined in the main study gave quite different results and we will go through them one by one, but the focus is going to be on the valence since that factor gave us the clearest results.

5.1 Valence

We can clearly see that there is a correlation between the facial static expressions expressed valence and the perceived full character expression (full body + face). The blurred faces and hands did not give us results that are clear enough to make any conclusions so the focus will be on the facial expression without blur (see figure 11) but they do however correlate to the findings in the prestudy (see figure 9). It was interesting to see that the body motions were perceived as more ambiguous when shown amongst animations where the facial expression was distinct than it was when shown only with blurred faces and hands as it was in the prestudy.

There is also a correlation between the perceived valence of the facial static expression in combination with the perceived valence of the body motion and the results of the main studies perceived valence. They both have the same staircase formation in the graph and the only noticeable difference is that the higher the values of the perceived valence that the body motion had, the higher the overall values of the perceived valence in the main study is while remaining more or less the same intervals in the staircase.

5.2 Other factors

5.2.1 Trustworthiness

For the trustworthiness the blurred faces give some interesting fact. The trustworthiness does not change between these body motions while the face and hands are blurred. It would be interesting to continue researching and checking to see if the trustworthiness ever changes between different body motions while the face and hands are blurred. Otherwise we might have found a tendency that trustworthiness is more likely to be judged by facial expressions rather than body motions.

As for the rest of the faces we can see that the facial expressions change the perceived trustworthiness in the same way as it changes the valence, but with less difference between the bars in the diagram. In the study conducted by Oosterhof and Todorov [16] it is shown that valence and trustworthiness have a relation in the way that trustworthy faces expressing positive valence is perceived as more positive than untrustworthy faces expressing positive valence. The findings in this study seems to add to that research as different valences are perceived as different levels of trustworthiness and that could explain why that study got the results it got.

5.2.2 Relationship

The data collected for the relationship between the characters is inconsistent and no clear correlations can be made. However, you can see a small tendency to a trend like the "staircase" in the two other factors discussed.

The only clear data is that a positive face makes for a perceived closer relationship between the characters. This might be because a smile often shows affection and since they are all smiling they seem to be happy about the situation and since they are standing close that might be perceived as a closer relationship. However, the rest of the data is inconsistent which raises the question about what shows that there is a distant relationship in a group. For future work it would be interesting to try nervous and uneasy facial expressions which would have been annotated beforehand.

5.2.3 Dominance

Dominance have the same problem with inconsistency as relationship but it has a much clearer "staircase" formation if you do not think about the blurred face and hands. But the standard deviation is really big on most of them so it feels as if the data cannot be trusted.

5.3 In relation to other studies

The most comparable study is the one conducted by [1] since their hypothesis was that the body motion would help determining the valence of ambiguous facial expressions. They concluded that the body indeed affected the perception of the emotion the person even if the facial expression was the same. The difference between this study and that one is that this one used moving images whilst that one used static imagery. This makes for a new interesting question, does the body and facial expression affect the perception of emotions differently if the body is moving or posing statically? Another difference is that they used grayscale images while this study used full color, that might influence the participants as well. The background has an impact as has been proven before and in this study white background was used and that could also be a factor that changes a lot.

In relation to the study conducted by [14], this study also worked with congruent and incongruent facial-body relations. But since they proved that a distinction was made after 120 milliseconds whether the relation was congruent or not we can assume that these stimuli are perceived as either congruent or incongruent since they were shown for approximately 10000 milliseconds. That raises an interesting question regarding what modality was analyzed by the participants. To determine that an eye gaze-tracker would have to be used.

5.4 Limitations and considerations

The size of the characters might be a limitation since they were not that big on the participants' retina. Each participant was sitting approximately 60 cm from the screen and the characters filled up about half the 24" screen in total, that makes for quite a small representation on the retina. We do not know if the size of the stimuli affect which modalities are significant while analyzing the emotion of the virtual characters.

A questionnaire was filled in by each participant and a question that should have been stated would be if the person was not neurotypically developed since we now have no information which were and which were not and that could possibly affect the results since people with e.g. Autism Spectrum Disorder (ASD) have troubles perceiving emotions in the same way as Neurotypically Developed Adolescents (NDA). We know that one of the participants is diagnosed with Asperger and therefore that participant's responses are not used in this study.

6. APPLICATIONS

The findings in this could be used as a base for a game that is in the making right now, called *Emotions with friends*. As part of the master thesis a demo of this game was made that would let children in a school environment learn more about emotions and how others perceive them (see figure 15). A study should be conducted to see whether or not this could be used for the purpose it was built for.

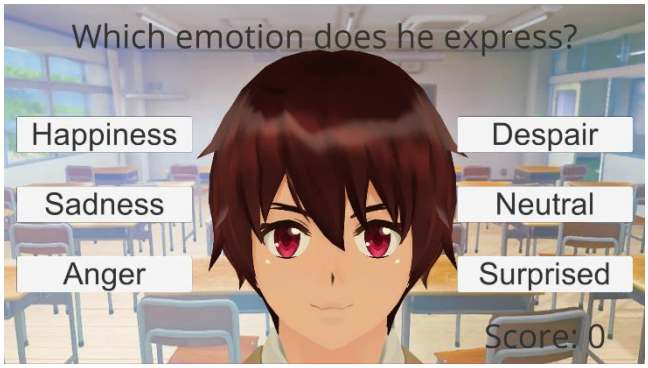

Figure 15 - Screenshot of an earlier demo of the game

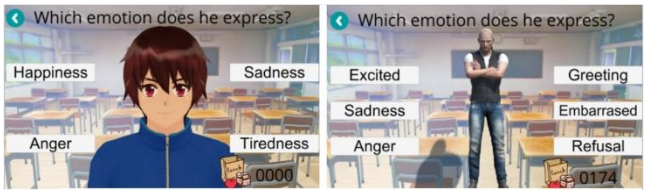

Figure 16 - Screenshot displaying both types of scenarios

The idea is to have two scenarios of the game (see figure 16), one that works with facial animations only and one which works with full body animations much like the ones tested in this study.

The game would be used in a school environment were the teacher would be the host of an online session. Then all the children would be presented with a stimuli and label that stimuli with a certain emotion. When all the children are done they would get scores graded on how many others chose the same label. This data could be collected and used in an open class discussion about the perception of emotions and why people perceive certain emotions differently.

If this game were to be spread across the globe it would be possible to collect a lot of data on how emotions are perceived differently depending on cultural backgrounds.

The data validated in this study could also be used to set up a set of games for children with ASD as to help with rehabilitation of the diagnose and let them better understand social situation and social cues in real life. There are games in the making with that purpose as well [13] and it is proven to have a positive effect [6]

7. CONCLUSION

It is clear that the facial expression has a big impact on the perceived valence of a group of virtual agents. Both ambiguous and distinct body motions were affected and congruent and incongruent facial-body relations all showed that the facial expression altered the perception of the overall perceived valence of the group. Whether it is the salient modality or just a part of the overall perception is left to be tested. But it can be concluded that the face was dominant in the perception of emotions in this study. Trustworthiness also gets affected by the facial expression, there is even a trend that might show that body motions does not affect the perceived trustworthiness of a virtual agent.

Relationship and dominance is less clear and might not actually be affected at all due to random responses. The only factor that might be concluded is that a positive face gives the impression of a closer relationship than a neutral or negative face.

7.1 Future work and improvements

All of this information is based on the assumption that virtual agents are perceived in the same way as human beings. To validate this one would have to do a study with an actor expressing a set of emotions, motion capture this. Use the motion captured data to animate a virtual agent that resembles the actor and do the same test as was conducted in this study. After that the recorded data of the actor would be used and show it with no post processing to see if there is any difference in the perception of expressions between virtual agents and human beings.

8. ACKNOWLEDGEMENTS

Thanks to Niklas Blomqvist who has helped a lot during this study. My knowledge of shaders, screens and projectors in combination with your knowledge of blend shapes and experience in Tobii Studio Pro made for a great team and I believe that both of our studies have won a lot from this cooperation.

I would also like to thank Björn Thuresson and Christopher Peters for lending us hardware, software licenses and facilities in which we could conduct our experiments.

9. REFERENCES

1. Hillel Aviezer, Yaacov Trope, and Alexander Todorov. 2012. Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science (New York, N.Y.)* 338, 6111: 1225–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23197536
2. Aryel Beck, Lola Cañamero, Antoine Hiolle, et al. 2013. Interpretation of Emotional Body Language Displayed by a Humanoid Robot: A Case Study with Children. *International Journal of Social Robotics* 5, 3: 325–334. http://doi.org/10.1007/s12369-013-0193-z
3. Miguel Ramos Carretero, Adam Qureshi, and Christopher Peters. 2014. Evaluating the perception of group emotion from full body movements in the context of virtual crowds. *Proceedings of the ACM Symposium on Applied Perception - SAP '14*, ACM Press, 7–14. http://doi.org/10.1145/2628257.2628266
4. James M. Carroll and James A. Russell. 1996. Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology* 70, 2: 205–218. http://doi.org/10.1037/0022-3514.70.2.205
5. Eirini Christinaki, Nikolas Vidakis, and Georgios Triantafyllidis. 2013. Facial expression recognition teaching to preschoolers with autism. *Proceedings of the 6th Balkan Conference in Informatics on - BCI '13*, ACM Press, 141. http://doi.org/10.1145/2490257.2490262
6. Kevin Conallen and Phil Reed. 2016. A teaching procedure to help children with autistic spectrum disorder to label emotions. *Research in Autism Spectrum Disorders* 23: 63–72. http://doi.org/10.1016/j.rasd.2015.11.006
7. Sofia Dahl and Anders Friberg. 2007. Visual Perception of Expressiveness in Musicians' Body Movements. *Music Perception: An Interdisciplinary Journal* 24, 5: 433–454. http://doi.org/10.1525/mp.2007.24.5.433
8. Cathy Ennis, Ludovic Hoyet, Arjan Egges, and Rachel McDonnell. 2013. Emotion Capture. *Proceedings of Motion on Games - MIG '13*, ACM Press, 53–60. http://doi.org/10.1145/2522628.2522633
9. Cathy Ennis, Rachel Mcdonnell, and Carol O 'sullivan. 2010. Seeing is Believing: Body Motion Dominates in Multisensory Conversations. *ACM Transactions on Graphics* 29, 4: 91:1–91:9.
10. Noemí Marta Fuentes García, Francisco Luis Gutiérrez Vela, Patricia Paderewski Rodríguez, Rafael López Arcos, and Natalia Padilla Zea. 2015. Teaching Emotions to Children by Using Video Games. *Proceedings of the XVI International Conference on Human Computer Interaction - Interacción '15*, ACM Press, 1–3. http://doi.org/10.1145/2829875.2829888
11. Mariska Esther Kret and Beatrice de Gelder. 2010. Social context influences recognition of bodily expressions. *Experimental brain research* 203, 1: 169–80. http://doi.org/10.1007/s00221-010-2220-8
12. Silvia Lovato and Anne Marie Piper. 2015. "Siri, is this you?" *Proceedings of the 14th International Conference on Interaction Design and Children - IDC '15*, ACM Press, 335–338. http://doi.org/10.1145/2771839.2771910
13. Laura Malinverni, Joan Mora-Guiard, Vanesa Padillo, Lilia Valero, Amaia Hervás, and Narcis Pares. 2016. An inclusive design approach for developing video games for children with Autism Spectrum Disorder. *Computers in Human Behavior*. http://doi.org/10.1016/j.chb.2016.01.018
14. Hanneke K M Meeren, Corné C R J van Heijnsbergen, and Beatrice de Gelder. 2005. Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America* 102, 45: 16518–16523.
15. N. N. Oosterhof and A. Todorov. 2008. The functional basis of face evaluation. *Proceedings of the National Academy of Sciences* 105, 32: 11087–11092. http://doi.org/10.1073/pnas.0805664105
16. Nikolaas N. Oosterhof and Alexander Todorov. 2009. Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion* 9, 1: 128–133. http://doi.org/10.1037/a0014520