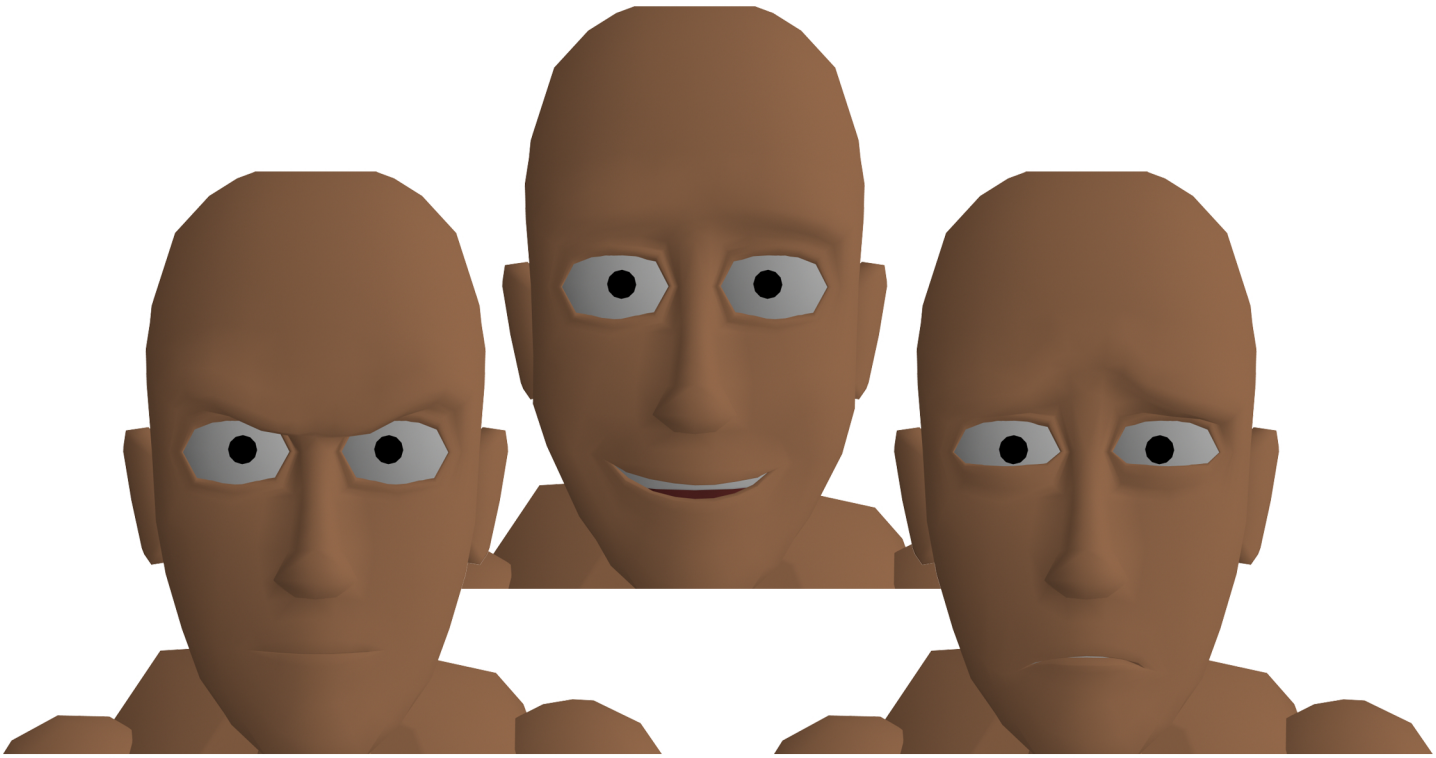
Principles of Facial Animation in the Expression of Emotions in Virtual Agents



Mikkel Lundbæk Jensen

Aalborg University Copenhagen
24-05-2012

Aalborg University Copenhagen

Semester: 10th semester

Title: Principles of Facial Animation in the Expression of Emotions in Virtual Agents

Project Period: 1st February 2012 – 24th May 2012
Summer 2012

Semester Theme: Master Thesis

Supervisor(s): Luis Emilio Bruni

Members:

Mikkel Lundbæk Jensen

Copies: 3
Pages: 53
Finished: 24-05-2012

Aalborg University Copenhagen
Lautrupvang 4B,
2750 Ballerup, Denmark

Semester Coordinator: Stefania Serafin

Secretary: Lisbeth Nykjær

Abstract:

This thesis investigates the relation between facial expressions and user experience with virtual agents.
One video were created from a choice of emotion, based on the six basic emotions, and another video with the purpose of exaggerate it, for a more Disney like expression. The chosen emotion was happiness, and was based on the features around the lower face region.

To evaluate how facial expression in virtual agents affect the user experience, a test to compare the two version were conducted, and get the test participants reactions to the sequences. A total of 31 people participated in the test.

The results showed that the majority were able to recognize the happy emotion portrayed in the character, but there was a tendency that they did not find the emotion on the character effective or convincing. There was a hint of the facial expression and body language of a virtual character, can have an influence on the user experience.

Preface

This thesis is written for the project “Principles of Facial Animation in the Expressions of Emotions in Virtual Agents” as a part of the work on the 10th semester study in Medialogy at Aalborg University, Copenhagen.

The purpose of this thesis is to see how emotions are affecting the user experience, which can be applied to learning applications to create better learning environments.

It is expected that the reader has knowledge of the terms connected to computer animation and to some extent emotions.

The thesis is build-up of a pre-analysis and analysis, which then will be used to create a model to work from, the so called Level of Resolution.

On the back of the report, one cd is attached together with this thesis. The cd contains the thesis in two formats, word and a PDF. The results from every test that was conducted as well as the two videos that was created for the project are also on the cd.

Acknowledgements

The author of this report would like to thank the people that have followed this project and for their assistance, participation and support during the project work. For reading through the report and come with feedback and editing, which have been very helpful.

Table of Contents

1.0 Motivation

The use of animations are becoming more common today in the different fields of our daily lives, in advertisements we see animated commercials, fully animated movies, even motion capture incorporated into real footage and of course animation in games. Many of these animations are having characters performing as either real-life looking characters or having human characteristics. The characteristics that I think are the ones really standing out are the various facial expressions of the different characters. The facial expressions is something that everyone is familiar with, as they are looking at other people's faces every day, and are seeing many different expressions (Terzopoulos & Mones-Hattal, 1997).

The animation of the facial expressions is an interesting part to look at, especially because the virtual agents can look very static in their respective environment, if the animation is not sufficient enough. This means that if the agents do not look alive, the user might lose the interest in looking at it for longer periods. Furthermore how are they affecting the users, when the user is looking at it, and how do they react to the visual representation of the character? If the environment is set up for learning, and the agent acts like a teacher, how would the emotional attitude of the agent portrayed by the agent affect the user?

1.1 Initial problem statement

How does a virtual agent's facial expression of emotions affect the user experience in an animated product?

2.0 Pre-Analysis

In this section I will be looking at the major topics that centre on facial animation, such as the expressions and emotions and virtual agents, which will be the general state of the art and background forming this project.

2.1 Virtual Agents

Virtual agents can be anything from a character in a game, to the character performing in a movie or the character present in an interactive application. The role of the virtual agent is often seen in cases where the agent is there to help you understand things better or for pure entertainment purposes.

A definition by Jack Krypansky (Agtivity.com, 2005)

> ***An Intelligent Virtual Agent (IVA) is a graphical representation in an interactive virtual environment (2D or preferably 3D).***

Where he also expand it, by saying that the intelligent virtual agent interacts autonomously and "intelligently" with the virtual environment, other IVAs, and with human users, and may act on behalf of another entity such as a human user.

Techtarget's definition of Virtual Agent (Techtarget.com, 2003);

> ***In customer relationship management, a virtual agent is a chatterbot program that serves as an online customer service representative for an organization.***

Mark Gaydos, vice president of VirtuOz (Chatbots.org, 2010);

> ***The ability to interact with a customer and, in a sense, understand what the customer wants to do and guide them to an answer to their question.***

David Lloyd, CEO of IntelliResponse (Chatbots.org, 2010);

> ***It is a way of interacting and delivering customer self-service.***

Ashutosh Roy, Chairman and CEO of eGain (Chatbots.org, 2010);

> ***A virtual agent is a human-like bot that understands natural language interactions and guides a visitor through a conversational interface using text and/or speech input and output.***

As can be seen from the different definitions of a virtual agent, they mostly focus on the fact that the virtual agent, acts as a chatbot figure, which is helping the user through different topics or even help them to purchase something online. Only Jack Krypansky from Agtivity defines the virtual agent as something that is a graphical representation in an interactive virtual environment, meaning that the virtual agent can do much more than just chat with the customers. Krypansky's definition of a virtual agent is also the most general definition, as he is taking into consideration that even a user controlled character in e.g. a game, is a virtual agent for another user, which someone else is communicating through.

The characters we see in movies are virtual agents as well, even though they are not in directly contact with the user, they are showing emotions and telling stories, which the user follows and either sympathize to or do not like. In animated movies it is through storytelling that the characters are giving us the experience. Disney is specially focused on the realistic approach to their movies and showing emotions with their characters, as it is the emotions that are giving the characters life and make them more believable (Bates, 1994).

With virtual agents in applications, the experience is gained through how good they are at helping the users with their problems and how user-friendly they are to the average user.

Virtual agents are also named in many different ways, but still covering the same concept. To name a few there is Intelligent Virtual Agents, Embodied Conversational Agents and Robots. What all the different names for virtual agents have in common, are that they are basically describing the same thing. This can sometimes make it confusing for people to figure out what they mean by these names, e.g. Embodied Conversational Agents are just describing an agent that appears as a virtual body, and is able to communicate with the users. It can be more graphical oriented than the Intelligent Virtual Agent, as the name is not telling you what kind of agent it is, so it can simply be a chat box that appears, and when the user is asking something, the agent replies with text only. Robots on the other hand, are the physical product of a virtual agent, the concept is still the same as in the definitions, but it also has the opportunity to do much more now, as it can walk around and do things just like humans.

A Classification of applications the virtual agent can be seen and their roles (Dautenhahn, et al., 2002);

- Socially Intelligent Robots
- Interactive Therapeutic Agent System
- Interactive Education and Training
- Socially Intelligent Agents in Games and Entertainment
- Social Agents in E-Commerce

A virtual agent that is a part of an environment about enhancing learning, Jie Yan and Ruth Agada's (Yan & Agada, 2010) experimental results showed, that a virtual agent that acted in a natural sense, with head movements and appropriate facial expressions, would get a much more positive experience compared to an agent that lacked these behaviours. In this research they only focused on the head movements and the facial expressions of the agent, and disregarding the body language, by the use of a virtual head only. In relation to real life teachers, this have the same effect, as if the teacher is not really engaged in what he is trying to convey, then the user experience would be more negative as well, as the teacher would not be very interesting to listen to for a longer period, and therefore get a lower learning experience, or none at all. Their research also showed that the students tried to impress the agent, as the agent always tries to motivate the student, by giving small engaged and motivating responses to make them try to do even better next time. Having a more life-like virtual agent can prove to be a good way of how to involve the users more, in every situation that the virtual agent can be implemented in.

To get the believable virtual agents, the agents must also resemble the human the best way possible in both movements and general appearance. Joseph Bates (Bates, 1994) had three characters, which had some behavioural mechanism programmed, and he realized that people were attracted to specially one of the characters, as due to a programming error, the character would randomly do something out of place, and people reacted to that and found it interesting. Chuck Jones found out and also stated that, it is *"the oddity, the quirk, that gives personality to a character, and it is personality that gives life."* (Bates, 1994) Same thing was said by two of the Disney's "Nine Old Men" [1], Frank Thomas and Ollie Johnston (Thomas & Johnston, 1981)*"From the earliest days, it has been the portrayal of emotions that has given the Disney characters the illusion of life."*

Jonas Beskow (Beskow, 1997) who studies communication came up with an architecture of how to design and animate characters of lip-synchronized synthetic speech, for use in spoken dialogue systems. Having an agent that very much resembles a human being also sets high requirements for the animation of the agent, because communication is multimodal: information is conveyed with body language, facial expressions, gaze, intonation,

[1] The Nine Old Men refers to the first nine animators at the Walt Disney Company, who created some of Disney's most famous works. They also refined the 12 basic principles of animation.

speaking style and are all important factors in everyday communication. He (Beskow, 1997) used the method of surface deformation, as this method is a fairly simple and efficient approach to facial animation. Surface deformation is also good to use on agents in interactive applications, as the method is a low computational, rather than have to compute the skeleton and each of the bone's location, rotation etc. for every frame in the sequence, when it have to be rendered out.

A big problem with virtual agents and their general appearances can be the uncanny valley. The uncanny valley is the relationship between perceived human-likeness and perceived familiarity of the characters. Tinwell (Tinwell, et al., 2011) researched in the relationship between facial expressions and the uncanny valley. Especially the upper forehead of the character, they found from a previous study, that the relationship between the upper face region and the uncanny in virtual characters had a significant importance on the graph, and the users' experience of that character.

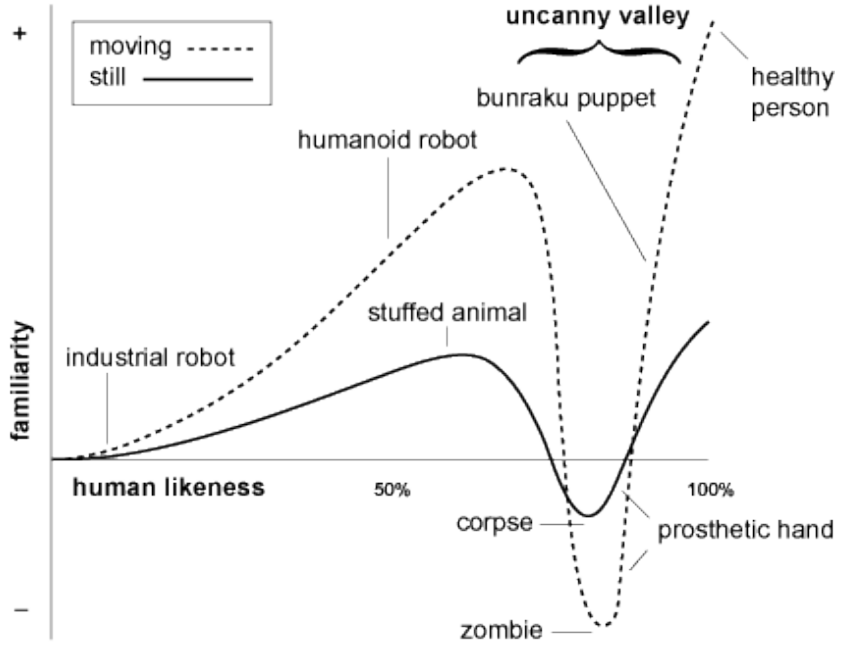


Figure 1 Masihiro Mori's "The Uncanny Valley", between the human likeness and familiarity

As the character are getting closer to look like a real person, there is the point at around 80% human likeness, following the moving line(dotted) in figure 1, where it looks real, but then again, it does not feel like it is real, and thus the familiarity is falling. On the graph there are two peaks for when the character have the highest likeness, at around 80% before it drops down, and then after the uncanny valley drop, where it leads up to the healthy person. As the technology have made it possible to increase the visual realism in games and movies, and almost quasi-realistic, human-like characters in games, designers are trying to aim for the second peak, as realism is believed to improve player experience and sense of immersion.

An example of a movie using motion capture throughout the movie, and based their acting on the motion capture data, and therefore realistic motions on 3D characters, is the movie Beowulf (Beowulf, 2007). Beowulf is purely made up of motion capture, and for the actors it meant that they were free to perform their scenes and actions throughout the shots, as there were no limitations of waiting for the camera or setting up the lights for a new scene. Everything was recorded digitally, but the end result was less effective though. Several times the characters looked like they were made of plastic or their appearance lacked "life".

Figure 2 A picture from the movie Beowulf © 2007

As can be seen in figure 2, the main character in the movie, is standing with his arms out in the particular sequence. When you look at the character, it just seems that something is wrong; it does look a bit like the character is made of plastic or at least not a realistic looking skin. This is also well summed up in a review of the movie (Gallagher, 2007):

> ***Perhaps we have spent so much time looking at our fellow humans that we can detect a problem with something as subtle as the physics of a muscle contracting, just a fraction of a second before it pulls the lips into a smile.***

Virtual agents often have a different appearance, dependent on the use of the application. As mentioned a virtual agent for some customer service can be a simple chat box, or it could be a picture of a character (sometimes a real person's picture) next to the chat box (see figure 3 A), representing the visual part of who the user is talking with.

Next is an example of a virtual agent in a movie (see figure 3 B), from the movie Toy Story (Toy Story, 1995), which also tried to have as realistic movements as possible, but fell into the uncanny valley, as the characters were trying to behave like humans, but looked too plastic-like, as also can be seen in the picture, that the light on Woody's nose and next to the eye is making it clear that they are only plastic toys.

The last example (Figure 3 C) is from the game Gears of Wars (Epic Games, 2006), showing two characters talking to each other. As the picture is showing, it is possible to see a lot of details on the characters, the wrinkles around the eyes and the mouth looking natural, and even the characters are looking more realistic, texture wise, but not so much proportional wise in this particular game, as the characters are supposed to look like some super soldiers.

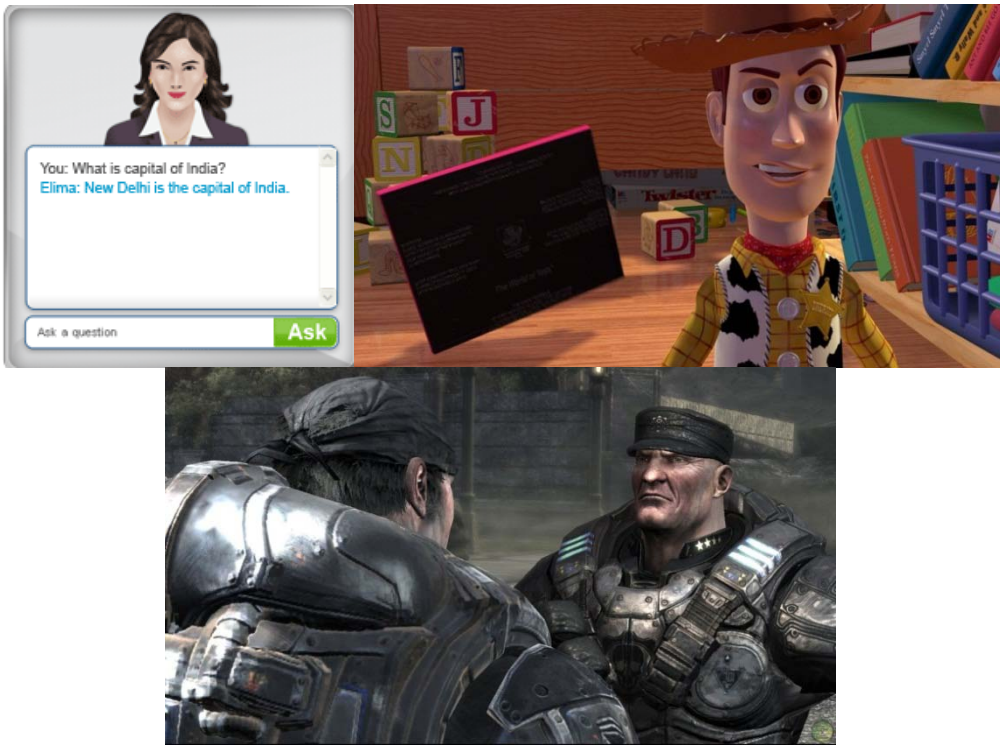


Figure 3 A: Example of a chatbot with a character acting as the agent. 2 B: Screenshot from the movie Toy Story © 1995. 2 C: Two characters in the game Gears of War © 2006.

2.2 Emotions

To get an idea of what kind of emotions there are, it is necessary to take a look at the research in the field of emotions. Emotions are a very complex field of study and have been researched for many years already, Ekman first found that the humans share six universal emotions, also referred to as the six basic emotions, which are cross cultural, and his results showed that if different cultures were allowed to use their own words in selecting emotion terms that fit the facial expression that was shown, they would to a great extent be the same (Ekman, 1971) (Ekman & Friesen, 1971) (Ekman, 1993).

Ekman's six basic emotions are the following;

- Happiness
- Sadness
- Anger
- Fear
- Surprise
- Disgust

To keep it simple and not go too much into the field of emotions in this project, it will be limited to the basic emotions, thus not go into the secondary emotions and the complexity of the relationships between these. Reason for this, is to keep focus on the emotions that are portrayed, and not the complexity behind the emotions.

To characterize how the six basic emotions appear on humans, Ekman also created a list that could be used to identify the different facial expressions, how the facial area would behave for each of the six basic emotions.

Emotion	Brows-Forehead	Eyes-Lids	Lower Face
Surprise	Raised curved eyebrows; long horizontal forehead wrinkles	Wide opened eyes with schlera[2] showing above and often below the iris; signs of skin stretched above the eyelids and to a lesser extent below	Dropped-open mouth; no stretch or tension in the corners of the lips, but lips parted; opening of the mouth may vary
Fear	Raised and drawn together brows; flattened raised appearance rather than curved; short horizontal and/or short vertical forehead wrinkles	Eyes opened, tension apparent in lower lids, which are raised more than in surprise; schlera may show above but not below iris; hard stare quality	Mouth corners drawn back, but not up or down; lips stretched; mouth may or may not be open
Anger	Brows pulled down and inward, appear to thrust forward; strong vertical, sometimes curved forehead wrinkles centered above the eyes	No schlera shows in eyes; upper lids appear lowered, tense and squared; lower lids also tensed and raised, may produce an arched appearance under eye; lid tightening may be sufficient to appear squinting	Either the lips tightly pressed together or an open, squared mouth with lips raised and/or forward; teeth may or may not show

[2] Sclera means the white area in the eyes.

Emotion	Brows-Forehead	Eyes-Lids	Lower Face
Disgust	Brows drawn down but not together; short vertical creases may be shown in forehead and nose; horizontal and/or vertical wrinkles on bridge of nose and sides of upper nose	Lower eyelids pushed up and raised, but not tensed	Deep nasolabial fold[3] and raising of cheeks; mouth either open with upper lip raised and lower lip forward and/or out, or closed with upper lip pushed up by raised lower lip; tounge may be visible forward in mouth near the lips, or closed with outer corners pulled slightly down
Sadness	Brows drawn together with inner corners raised and outer corners lowered or level, or brows drawn down in the middle and slightly raised at inner corners; forehead shows small horizontal or lateral curved and short vertical wrinkles in center area, or shows bulge of muscular contraction above center of brow area	Eyes either glazed, with drooping upper lids and lax lower lids, or upper lids are tense and pulled up at inner corner, down at outer corner with or without lower lids tensed; eyes may be looking downward or eyes may show tears	Mouth either open with partially stretched trembling lips, or closed with outer corners pulled slightly down
Happiness	No distinctive brow-forehead appearance	Eyes may be relaxed or neutral in appearance, or lower lids may be pushed up by lower face action, bagging the lower lids and causing eyes to be narrowed; crow feet apparent, reaching from outer corner of eyes toward the hairline	Outer corners of lips raised, usually also drawn back; may or may not have pronounced nasolabial fold; may or may not have opening of lips and appearance of teeth

Table 1 Ekman's identifications of the six basic emotions (Ekman, 1971)

[3] Nasolabial fold means the smiling lines around the nose when you are smiling.

As can be seen in table 1, these are some key features that appear in the real emotions, how a human's face would appear when one of these emotions is in play. Each of the emotions are using either primarily one of the areas of the face or a combination of one or more to make certain facial expressions. This is especially important when it comes to choosing what emotion(s) that is going to be used in this project, and what area that should be focused on, when animating the facial expressions. This list can also to some extent help spot fake emotions, as there is typical some muscles that are not being used when faking an emotion e.g. be used to please another person, like the smiling at a gift that you did not really like.

Transferring emotions to a computer generated character is not just about showing a happy face, there are five factors that need to be considered, that is essential for the believability of the character. Bates (Bates, 1994) describes the first three of the five factors that Thomas and Johnston made on points to remember when animating emotions (Thomas & Johnston, 1981, p. 507)

1. *The emotional state of the character must be clearly defined.* The animator must know the states for every moment, in order to make the viewer able to interpret the emotion status of the character.
2. *The thought process reveals the feeling.* The audience should be able to see the emotions through how the character acts. The animator must make it seem to the viewer, that the character is influenced by its emotional state, a simple move or gestures.
3. *Be alert to use of cutting and camera in helping to accentuate the emotion.* There must be no unnecessary cutting or camera movements to show the emotion as it in the worst case, can break the viewer from getting the emotional state of the character.

   The last two factors are, and completing the five factors:
4. *Ask yourself constantly:* what are you trying to say here, what do I really want to show, how do I want the audience to react?
5. *Use the element of time wisely*: to establish the emotion of the character, to convey it to viewers, to let them savour the situation.

The last two factors are as stated, more relevant for asking yourself, how does this look like and what do I want to show the audience. The same is for the time; the action needs to be built up, as there must be a reason for the character to do things.

Taking these five factors into consideration, it is not only the face of the character that must be showing a feeling, but the whole body must show the emotional status of the character as well.

2.2.1 Secondary Emotions

Secondary emotions will not be considered in this project, but will shortly go into the complex field of secondary emotions and what is happening when having these.

Matthew Tull (Tull, 2009) is working with Post Traumatic Stress Disorder and he defines secondary emotions as:

> ***The emotional reactions we have to other emotions. For example, a person may feel ashamed as a result of becoming anxious or sad.***

Elaborating on this definition he explains, that the secondary emotions often are caused by the beliefs we have about experiencing certain emotions. In this case, if some people are being sad, they may think it is a sign of weakness or being sad may say something negative about them as a person. Therefore every time they have these emotions, they are thinking these thoughts and that triggers the secondary emotions that are not clearly visible, as they happen inside you. This is also a reason that the use of secondary emotions will not be used for this project. The secondary emotions are still quite important, but difficult to portray and work with, as it happens on a more personal level, and not a visual level such as the facial expressions.

2.1 Pre analysis sum-up

To keep the project simpler, it has been chosen to work with only the six basic emotions by Ekman, as the use of more complex emotions and secondary emotions would be too difficult to implement. The aim of this project is to investigate the relationship between the visual facial expressions in virtual agents and the users experience and their reaction to these.

Thus the choice of emotion that is going to be used for the test purpose will be limited to one of the six basic emotions. The emotion that will be the best to work with is either happiness or anger. These are the opposite of each other, but also two that are different in how the facial expressions are formed. Anger is primarily using the forehead, and a little bit with the mouth. Happiness is primarily using the mouth with only little forehead movements, minor movement around the eyes, which is caused by the upper movement from the mouth.

Virtual agents are the character that you interact with in either games or see in an animated product. To get them to look more realistic, adding odd behaviours as Bates (Bates, 1994) did it, though they did not do it intentional, but that a programming error led to that this character got more attention, as it stood out from the other characters. So a factor to consider for later use is the unexpected actions that the users would not expect to see from the character. As Jones (Bates, 1994), Thomas and Johnston (Thomas & Johnston, 1981) all talked about, it is the emotions that are giving the characters the illusion of life.

To animate the facial expressions, there are several methods as well. The animations can be programmed or through the use of motion capture, and the work that have been reviewed, have been with the use of these processes. Therefore I will try to manually animate the characters facial expressions, to see if the same kind of results can be achieved. The manual method can be done by either through the use of surface deformation or by the use of a thoroughly made skeleton rig, full of bones.

The simplest and most low computational method is the surface deformation, as the skeleton rig can quickly become very computational, because it have to compute the location of each of the bones, and its attributes, such as rotation, direction. With surface deformation, the computer would only have to compute the location of the changed surface once for each frame. Of course the character is still build up by a skeleton which have to be computed, but for facial animation, surface deformation makes it require less to compute.

So the main focus on this project will be to manually animate the facial expressions that are needed for this project, and create two scenarios of a realistic and a more exaggerated Disney version.

3.0 Analysis

In the analysis I will take a closer look at the different methods of how to animate characters, focusing around facial animation and what the possibilities are for this project. Then take a look at Affective Computing and how it is related to virtual agents.

3.1 Animation methods

In this section I will look at the methods that could be used for this project, as the aim is for a more realistic look and at the same time focusing on the facial expressions, methods such as stop motion, hand drawn, Claymation etc. will not be considered.

3.1.1 "E-motion" Capture

Motion capture is the newest kind of animation method that has been developed, as the computer industry grew larger, and the use of motion capture is increasingly being used in many situations. Commercials are one of the big categories for using motion capture, as many commercials are using other means of portraying their message than using regular humans on the screen, e.g. the Tulip commercial with the bear (and their other version with a crocodile) are both completely done with CGI (Computer Generated Images), the background is computer made. For the motions of the bear (and crocodile), the use of motion capture have been applied, with special focus on the facial expressions and the speech (Ghost VFX, 2009).
In their walkthrough of how they did it, it can be seen that they used a marker less setup, where they are tracking the areas they need directly from the computer, for these commercials, the speaking and the mouth movements were in focus. The great thing about the marker less setup is that it is easy to get to start shooting the sequences that is needed. The other method is with markers, where the actor has to wear several markers attached to the face or on the body.

The equipment used for motion capture can limit the space that are available to perform in, as a common setup for full body motion capture is using eight cameras, which should be able to capture all the different movements that the actor is performing, while recording. The performance between both systems is more or less identical, Katherine Steele et al. (Steele, et al., 2009) performed several tests on both procedures, and realized that the difference was minimal between using markers or marker less motion capture setups. So the method that would be the easiest one to use for both the producers and the actors would be the marker less setup, as they do not

need to spend time on setting up the markers and calibrate the system, but can go directly to shoot the sequences needed.

Facial motion capture is about capturing the external emotion of the characters, and therefore can be called, emotion capture. The procedure is generally the same with both the marker and marker less setup. Though the focus for the facial expressions are not as with the rest of the body, tracking the different joints of the body, but rather the special features on the face and muscles, such as the eye brows, cheeks, chin etc. The advantages of motion capture is that the data can quickly be recorded in a small studio, and the data can be transferred directly to any 3D character and the animation would then be applied to the character and can be used as where it is needed. These data can even be saved for later use in other projects with other characters too. A major drawback for motion capture is that if something needs to be changed either major movements or something needs to be exaggerated, it is very difficult to manipulate, as the control of the data is already set for the specific motion.

3.1.2 3D Animation

3D animation is more like the traditional animation style, hand drawn animation. The only exception is that the animator is moving the character in a virtual environment and gets more freedom to view the character from different angles and move the character around as he likes.

The advantages of 3D animation compared to motion capture, is that the animator gets more control of what is happening in the scene and what the characters are doing, and it can be changed quickly if it is needed.

A disadvantage is that the character animation can be very time consuming, as almost every single bone in the character have to be manually moved individually every time and then check if the movement look as it should in regards to the storyboard that is being followed.

An advantage of doing it all by hand, and not use motion capture or any pre-recorded data, is that the animators have the possibility of building up a character with certain character traits, which would be more difficult to achieve with motion capture. And also by doing it by hand, it is easier to exaggerate the different actions.

3.2 Affective Computing

Affective computing in general terms, means that the computer should be able to show and respond to emotions. As affect is another word meaning emotion, this whole field of research is a good starting point for how to improve virtual agents in the different applications.

Affective computing is a new field of study, and is still a young field according to Picard (Picard, 1995), the research results so far is only the tip of the iceberg, within the recognition and synthesis of facial expressions, as there are also a lot of physiological measurements that should be considered.

Whatis definition of Affective Computing (Techtarget.com, 2005);

> ***Affective computing is per definition human-computer interaction, where the device has the ability to detect and appropriately respond to its user's emotions and other stimuli.***

Oxford English Dictionary's definition of Affective Computing;

> ***A branch of computing concerned with the theory and construction of machines which can detect, respond to, and simulate human emotional states.***

Affective computing is centred around the computer acting like a human, as can be seen from Picard's four categories about affective computing that has focus on expression and recognition (Picard, 1995);

1. Computers that cannot express affect or perceive affect. Neither personal nor friendly
2. Develop computer voices with natural intonation and computer faces with natural expressions. In most cases seen as a virtual agent. This category cannot perceive affect but can express affect.
3. Enables computer to perceive the affective state and enabling it to adjust its response in ways that it can make it a better teacher and more useful assistant. This category cannot express affect, but can perceive it.
4. Maximizing the communication between human and computer, potential of providing truly "personal" and "user-friendly" computing. This category can both express and perceive affect.

The first category is basically the normal computers that are standing around in many homes and other places, such as institutional areas and companies. They are only being used as a working tool.

The second category, which is also the most relevant to this study, is the involvement of the computer having a voice and the ability to show natural expressions. Of course it is not the computer itself showing the emotions, but as stated a virtual agent in an application that the user is using.

The third category is the next step from the second category, making the computers able to pick different responses according to the emotional state of the user, to become a better teacher that is able to encourage the students, by not giving the same response to every student, but take into consideration of the emotional state. This is also here it begins to become unreal for many people, if they are afraid of computers that are being driven by emotions.

The last category is where the computer is both able to express and perceive emotions, and the purpose to maximize the communication between humans and computer. This category is about the aim of creating a computer that could walk among us or control buildings from a single place. A physical example would be the robots that are in the movie I, Robot (I, Robot, 2004), where they are walking around with normal people in the daily life, helping them at home or at work. A more non-physical example could be the computer HAL 9000 in the movie 2001: A Space Odyssey (A Space Odyssey, 1968), where the computer is not in a physical form, but acts as the artificial intelligence that is controlling everything in the spaceship, and is represented as a red television-camera eye on the equipment panels throughout the ship.

The social interaction or responsibilities between human and the computer is becoming more common, as there are many different applications that are using a virtual agent as the main character, which the user is following through a series of steps to complete either a game or an assistant that is helping you with some problems through a web chat or something similar.

To take the jump into the computer, the virtual agents are often seen as graphical embodiments, which are called Embodied Conversational Agents (ECA). These characters have human-like appearances and behaviours for the purpose of interacting with humans. (Peters, 2011) As already mentioned communication is multimodal, and therefore it is important to know how to coordinate the facial expression, speech, gestures and other modalities to make the ECA show appropriate emotions.

In order to believe in the agent, which is the general case for all graphical productions trying to sell a story, the agents must look believable and show expression of emotions in a consistent and natural look way across all modalities. (Martin, 2011)

Affective computing offers a wide range of applications, in the field of virtual agents, e.g. E-learning where the agent can see if the user is having difficulties with a topic, the agent can offer some expanded information or explanations about the topic and give examples.

Depending on the application that is being created and for what purpose, the agent can learn to recognize the user's posture, gesture and facial expressions. Then agent would then evaluate the situation and respond in a proper manner, or even adjust the application for the users need, and therefore enhance the user experience or productivity. The virtual agent would become a partner, rather than an agent telling you what to do.

4.0 Level of Resolution

In this section I will go through the previous chapters and create a model that will be used in connection with the practical part of the project. I will sum up the important things from the pre-analysis and the analysis, and discuss the different possibilities and trade-offs with the model.

Starting from the beginning, we have the virtual agent, which are being called in many different ways, and such as Embodied Conversational Agent, Intelligent Virtual Agents and Robots. Each of these names for a virtual agent is covering the same ideas, but is having a different opinion on the appearance of that agent. Robots are the physical form of a virtual agent, but are still considered an agent, because it is still only a computer in the form of a human. Embodied Conversational Agents that are the agents represented as a graphical entity, and therefore the user is able to see who they communicate with, and relate to it in another way, than with an ordinary Intelligent Virtual Agent, which is more like a computer responding in text only, and no graphical relationship between the user and the agent.

The different definitions of what a virtual agent actually is, is also heavily based on the intelligent Virtual Agent, that most companies are using an agent for customer support, to help their customers with their problems and guide them to solutions. These agents have some pre-defined answers, so if the customer is asking about a topic, the agent is looking at the category and finds the pre-defined area for that category.

One definition are more general, and taking more areas into consideration, such as it is a graphical representation of either an agent used in customer support, or even as the character in games or movies. Thereby widening the area of virtual agents and what they imply.

The interaction between the graphical agent and the non-graphical agent is better, as the user can see a reaction to any actions done and it is easier for the user to create a relationship to the agent. This is for the general definition of virtual agents, and with relationship I mean that the user can pay attention to how the agent reacts,

relate to its feelings, and relate to situations. This can happen in movies, games or in the graphical representation of the customer support.

Particularly in games, the interaction between user and agent, is limited to a certain degree, but the emotional relationship is much greater, especially if it is a character that have been used for a long time, and the user is really attached to that character. If the user have used a lot of work to build up a good character in a game, and suddenly have to stop playing that game, the user often feel a little disappointed or having mixed feelings of letting it go.

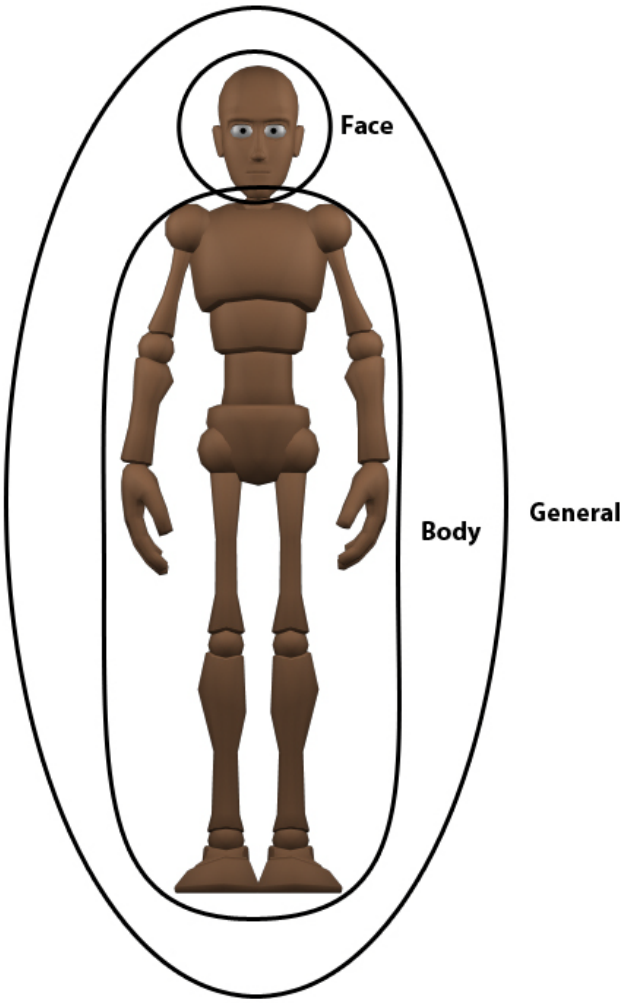


Figure 4 Model of the different areas to focus on, depending on the use.

In figure 4, we see the model of the character, which is also the same that is used for this project, and it shows the important areas to focus on, dependent on several factors. Some of these factors could be emotions, actions or the character as a whole entity.

It is important to focus on the face, when the character has to:

- Talk with others (conversations)
- React to something (emotional state)
- Direct attention to something
- Micro expressions

In games the face region is usually used in cut-scenes, where the characters are talking to each other, as the story develops between the different characters in the game. The focus on the face is to get the facial expressions as well as make the user able to see the lip movements whether they are trying to match the actual words or rather moving the lips to make it seems like it is following the words. Doing the facial expressions is also an attempt to establish a connection between user and the character.

In movies the face region is used when the character is in situations where the facial expressions is the most important to portray the emotions that are being felt, and often also let the viewer's sympathize with the character, as the characters are then trying to make the best of it. Here comes the emotional state of the character also into play, to establish the connection to the audience, and show that even though they are just fictional characters, they have emotions as well, and therefore the audience quickly relates to them.

Using the facial region to direct attention can be used simply by having the character looking out of the screen or away from the audience's view of the scene. An example could be a character running and looking back over his shoulder, and thereby create the sense that there is something important behind the character, making him run. Another example could be the character is coming towards the camera, just climbed a big hill or another big obstacle, and then stands on the top gazing over the area, with wide open eyes and an amazed looking facial expression.

Micro expressions is also where the face region is wanted to be in focus, as the micro expressions would otherwise be very difficult to perceive, if the camera is far away from the character and probably lose its meaning in the given context, because you did not see the expression that the character had at that moment. Micro expressions could be muscles moving on the face, the character giving another character a quick eye blink or the eye brows are raised in surprise or lowered in contempt. If the face is not in focus when these small, but relatively important actions are being performed, the audience might lose the meaning or context in the scene. If the camera was showing the whole body, the audience would not know that they should be looking at the face region of the character in order to get the full meaning of the scene.

Applications of use:

- Chatbots

Chatbots are usually a static picture of a 3D character, which represents the given company, it works as a customer service agent which helps and guides the customer in the right direction regarding their problems. Even though the usual chatbot is a static picture, it creates a more welcoming feeling knowing that you are talking to "someone" rather than something, in this case a computer with pre-programmed answers, based on keywords.

It is important to focus on the body, when the character has to:

- Direct attention to something
- If special parts of the body is important for a particular scene

Directing attention to something can be used in several cases, and can be used in all of the regions, but the uses for the body to direct attention in this case, could simply be that you see the character pointing towards something either in the scene or outside the scene.

Special parts of the body, is also a possible use. If the characters are dancing and the viewers need to be aware of small actions happening either with the feet or the hands, and therefore the individual body part can be in more focus than the rest of the body.

Applications of use:

- Games

Some action games have some built in action sequences if you do a special combo or push the right buttons at the right time. E.g. in the newest game of The Elder Scrolls: Skyrim (Studios, 2011), when you are fighting against a dragon or another enemy, suddenly the view changes from your current perspective to a more action camera based view, where the character is automatically attacking the enemy with a series of hit, while you just watch the action playing out. It usually happens when the enemy is about to die, and you hit the right moment. This feature is also helping making a game more experiencing or fun to play, as you would normally just play it for the story, but when this happens the first time, you start to see if it is possible to trigger other sequences. The story is usually the main focus while playing and will continue to be through the game play, but more and more games are letting the players go beyond the storyline and let them explore the world as well, and let them do whatever they want and finish the storyline along the way.

It is important to focus on the general region, the whole character, when the character has to:

- Perform body language and facial expression
- Direct attention to something

Having focus on the whole character is the normal way of watching characters. Scenes like these are usually when the whole setting of the scene is important, rather than the individually character traits, movements or facial expressions, and therefore can show the scene from an outside perspective. It could even be that the setting was the most important, and therefore you are seeing the environment more than the actual characters in the story.

Applications of use:

- Movies
- Games to some extent

Movies are typically the application where the whole characters are used the most. For a lot of storylines in movies, the environment is very important too, to establish the location and time that people can relate to, either by having heard of the place or actually been there.

A movie that is heavily based on the environment is the movie Rio (Rio, 2011). In a sequence you see the two main characters, two blue parrots, being followed by another parrot. During the chase, they are sliding on a piece of metal down the roofs of the suburban part of Rio, which ends with the parrot following them hits the transformer powering the city with electricity, and the lights are going out all over the town, and on the stadium where an important soccer match was being played.

4.1 Choice of Emotion

The choice of emotion that will be used, will be based on the possibility of how much it can be exaggerated, for the purpose of comparing realistic to a more Disney approach, with the concept of *"if it needs to be sad, make it sadder"* (Thomas & Johnston, 1981, p. 65). To make it more simple to actually implement as well, I will look at how much of the area of the face is being involved in each emotion, and choose one that are centred around a certain area on the face, as facial expression can be a very complex thing to work with.

So bringing up the list of the six basic emotions again (Table 1), we see that the general actions for the emotions are;

Surprise – Raised brows, wide opened eyes and dropped-open mouth.

Fear – Raised brows, eyes opened and more raised lower lids, mouth corners drawn back, but not up or down.

Anger – Brows pulled inward, upper lids appear lowered and lower lids raised, lips are tightly pressed together.

Disgust – Brows drawn down but not together, lower eyelids pushed up and raised, mouth open with upper lip raised and lower lip forward, and deep nasolabial fold.

Sadness – Brows drawn together with inner corners raised and outer corners lowered, eyes glazed with drooping upper lids and lax lower lids, mouth either open with partially stretched or trembling lips or closed with outer corners pulled slightly down.

Happiness – No distinct action in forehead, eyes are relaxed lower lids may be pushed up by lower face action, outer corners of lips raised usually also drawn back.

A quick look on the list and it shows that the emotion that are centred on one area is Happiness, the lower face area, Happiness is only involving the upper face because of the action of the lips being raised, and thus pushes the cheeks upwards towards the eyes, and make it seem, that there is an action in the upper face area as well.

Another emotion that is relatively centred on one area is Anger, which is primarily using the upper face area, such as the eyes and brows being moved inwards and lowered, pulling the areas together. The small secondary action is that the lips are being pressed together, but no major movements as such in the lower parts of the face.

To keep the emotions relatively simple and easy to recognize when testing, an emotion that would be easy to replicate without involving too much of the face would be preferable. Ekman and Friesen (Ekman & Friesen, 1971) tested on adults from New Guinea, who had very little contact with people from the Western culture, whether to see if they could recognize western facial behaviours. They found that they generally had a problem with

separating the emotion Fear from Surprise, and even the people from the Western culture that were tested, did not do any better, even though the facial behaviours were to be found native for them to recognize.

Also to not risk trying to portray one emotion and cause the test participants to see another. For this experiment the emotion happiness will be used as a standpoint, which is centred on the lower face and the mouth region, and no upper face movements as such.

5.0 Experiment Design

To begin with I had to find everything that I needed to create the scene. As I want to focus on the visual aspect of the animation, I will find a 3D character that is already rigged and set up for immediate use. If I had to create my own character and rig, it would be too time consuming, and would also be a completely different project, about creating a character instead of utilize one.

The character and rig must be free to use, and therefore I know, that I will not be getting characters that are very realistic looking, as most of them are expensive. I have been using a few free ones for my own small projects that I will be looking at, to see if they also have the features for facial animation that I need. The first one is called AndyRig (Jdoubles, 2007), which I used for a couple of self-study projects as well, even though it look semi realistic in the figure, I did not find it very good to use.

Another free rig that I had been using before was the goonRig (Burgoon, 2010), created by Sean Burgoon, Digital Artist and Animator at Industrial Light and Magic. I found this rig better to use, both for controlling the character, but also the controls for the facial animation, and therefore was my preferable choice. Though there is a disadvantage about the rig, which I have to take into consideration when talking about test results, as the character does not resemble a human in the form of a figure. It is created by spheres and circular squares as body parts, forming the human body.

The goonRig is inspired by two other rigs, one called Norman Rig, which was created for the purpose of the use with students taking the Pixar classes at the Academy of Art University. The other rig called Bishop, which is a rig only available for students at the Animation Mentor education.

The goonRig is a complete setup with character, with the skeleton and rig attached as well, and therefore it is easier to focus on the visual, rather than the technical of creating a character and setting up a proper rig and make it work. Thus for this project, this rig will be used for the animations.

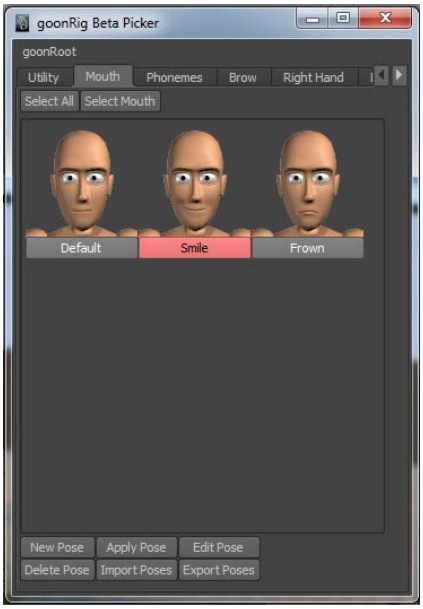


Figure 5 goonRig Picker Interface

With the goonRig followed a small interface (see figure 5) for quick selection of the different body parts, but also a small library of expressions that could be used for the character, one of them was for a happy, smiling character. The library is customizable, so it is possible to add more poses, both for facial expressions but also for the rest of the body. Which make the rig a great tool, as it can help the animator save time while animating.

5.1 Setting the scene

The small scene (see figure 6) where I placed the character, is built upon a scenario that would be easy to relate to and then the human's general ability to judge people only by the first look.

Figure 6 The scene setup

The scene is based on a school setup, but can be applied to anywhere, where the viewer would act as a student (or anything else), and see the teacher (or another person) coming into the classroom (or another place).

To begin with, the viewer cannot see the teachers face, as only the back of the teacher is visible, and therefore only base their judgement on the body language. This is to see if you can portray one emotional state and then when they see the face, getting surprised that his actual emotional state is completely different. To use this option best, I made the character look demotivated, by having him walk slower and shoulders hanging lower than usual, and even the head tilted down a bit, to make him walk slower I set that each step would take 20 frames, which means that he would walk like an elderly or tired person (Williams, 2001). This could sound like a description of sadness too, but in this case, it is the demotivated look, that I will try and portray. If the test participants then talks about the character being sad, it could be seen as a demotivated person as well.

5.2 First Iteration

Now I had the character and it looked happy (see figure 7) and was based primarily on the build in interface for a smiling face, but also according to Ekman's description of happiness. I needed to see if others also saw this as a happy character. A quick initial pre-test was conducted, showing the character to 10 persons and asking them to describe his emotional state. The pre-test showed that 6 out of 10 persons believed that the characters facial expression was a happy expression, and therefore I decided to continue with this expression and create the Disney version of it. (Appendix A)

To create the Disney version of the happy character (see figure 8), I first had to think about which principles that could be used to create a more Disney like character that was happy. To do that, I looked at the 12 basic principles of animation (Thomas & Johnston, 1981, pp. 47-70), and decided to use the exaggeration, as the point of having the Disney emotion is exaggerating the actions that are portrayed. I first started to exaggerate the smile a lot more, by making the outer corners of the lips coming much further back on the face. Then the cheeks were changed as well, making it more puff like, as the actions from the lips are pushing the cheeks up towards the eyes and making them come out.

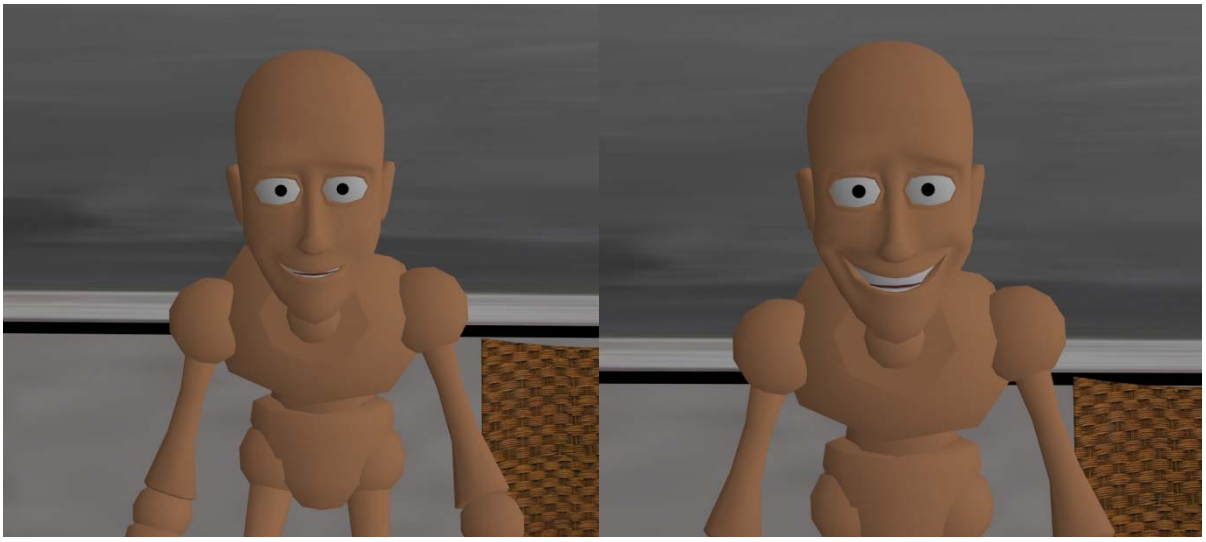


Figure 7 & 8 Left one showing the realistic version and the right the Disney version of the happy emotion

As the focus of this project is on the facial expression, the body language on the character in the Disney version will be unchanged. I believe that changing the body language, in this case exaggerate it more, could give some interesting results as well, but as already mentioned, the focus here is on the facial expression. The body language is a big part of the daily communication, but one of the first things that people base their judgement on, in a person, is the eyes, because when you talk with another person, you usually look into their eyes, to show interest in what the other person have to say.

5.3 Second Iteration

During what I thought would be my actual test (Appendix B), I noticed that several of the test participants commented on the quality of the video, but also specially on one thing on the character, the eyes looking creepy. Even though that from the first pre-test, the general opinion was of a happy character. Comments about the quality such as "*kinda surprised I'd say, kinda hard to see because of the quality*" and "*Can't really see it - it's too small...*" made it more obvious that it would not be preferable to continue, and therefore I decided to take the comments into consideration, and then my actual test became a second pre-test. I changed the settings for the quality of the rendered images to 1920x1080, which is the considered the best HD quality, to make sure that the quality of the video would be the best possible.

The comments about the eyes such as "*He is smiling (happy) but the eyes are making him look creepy*" and "*I was surprised that he was smiling, and thought he looked kind of creepy, since his eyes look sad but he is smiling, and the way he looks around, just smiling when seeing the students*". Comments like these, made it obvious that something needed to be done with the eyes of the character.

For the realistic version I started to change the appearance of the eyes, and this hopefully makes them less creepy, and more towards a happy face, with more neutral eyes (see figure 9).

To start with, the pupils were changed as I felt they stood out too much, so I made them a little smaller. The next thing that I had to look at, was the cheeks pushing the lower lids upwards, I puffed the cheeks a bit more, so that it looked like there were some more facial action from the smiling, and moved the cheeks upwards towards the lower lids, and raised the lower lids a little as well. I placed the upper lids back in its initial position, as I think it gave it a more relaxing or neutral look, rather than the brows seems to be lifted up as well, which should not be the case according to Ekman's description of the emotion happiness.

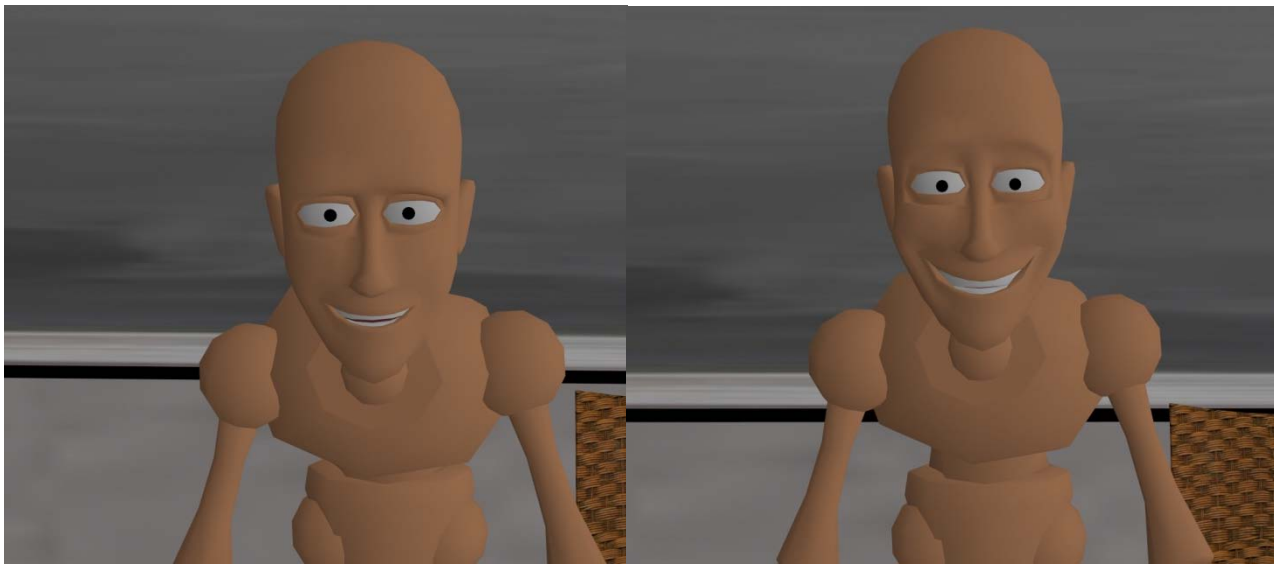


Figure 9 & 10 Realistic 2nd iteration on the left and Disney 2nd iteration on the right.

For the Disney version and the second iteration of this (see figure 10), I mostly changed a few things centred on the eyes, such as the lower lids being pushed up more by the cheeks, and cheeks being pushed up a bit further as well. Then I decided that since it is an exaggerated version and can be more Disney like, I looked at different Disney movies, and saw a tendency to raise the eyebrows. Specially one of their latest movies, Tangled [REF], where the characters seemed to have a happy emotion all over the face, and not only in the mouth region as opposed to Ekman. This is of course also possible, and to some extent accepted, as the characters are nowhere near realistic in the sense of photorealistic portrayal (they are after all only CGI), but still with realistic movements and expressions.

In the 12 principles of animation, exaggeration can be used for a lot of purposes; one of them can be to make a funny scene a bit more fun, as the little extra can make it look better, even though it might not be possible in the real world. Another thing it could be used for is to make a character do things that would normally not be

possible, balancing while almost lying on the ground and still get back up on your feet (which might also have the humoristic touch on the viewer).

5.4 Final Test

The test was conducted on the internet and the questionnaires were posted on several forums as well as on Facebook. On CGSociety, a forum for computer graphics, there were posted a description of the test and instructions of what to do, and then 2 links, one for each questionnaire. The link was also posted on social forums such as Facebook and Reddit where a single link was posted with a description of the test. The participants were directed to a page with 2 links, and then asked to choose one of them, then they were directed to the questionnaire of their choice and then followed the instructions in the questionnaire, and either watched Video A or Video B first. The quantitative and qualitative parts of the questionnaire are only based on the first video they see. When they get to see the second video, there is only one question that works as a comparison between the realistic and the Disney version.

6.0 Results

In this section I will look at the data that was received from the test. Check how the test participants answered according to their experience and eventually see if their answers might indicate a connection between expressions and user experience. The total test participants for this test were 31. 17 who watched Video A first and then Video B, and the 14 others watched Video B first then Video A.
As the test results are based on a visual context, such as a facial expression, each test participant may also have a different definition of what they think an emotion is and how it looks like; I will try anyway to see if there are tendencies based on the test participants answers and comments. All the iteration results and the final test results are all attached on the CD. The results for the final test can be found in Appendix C

The first question in the questionnaire was to see if the test participants could recognize the characters mood, based on his facial expression.

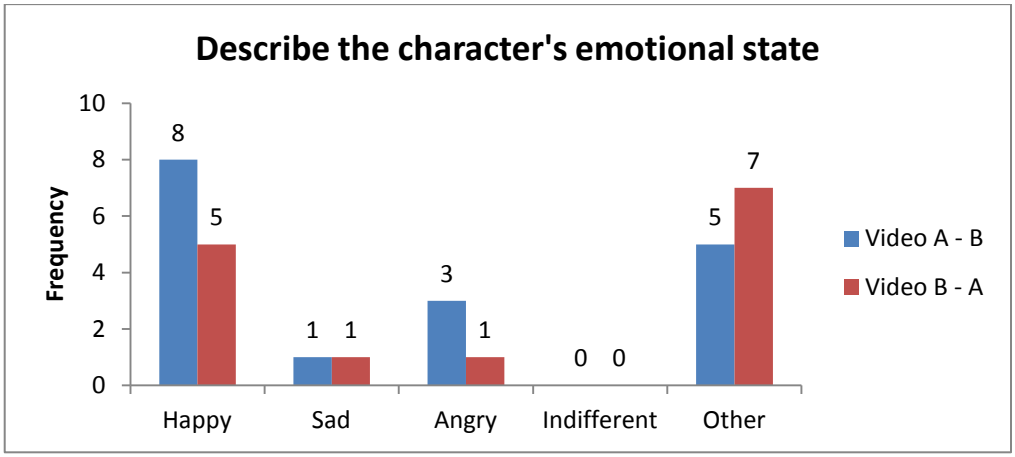


Figure 11 Description of the character's emotional state

As can be seen on figure 11, there were 2 groups for those watching Video A first, the biggest group of 8 saying that it looked happy and 5 who said something else in the category "Other", where some of those statements were emotions like: "Unstable", "Nervous", "Insane". What I can say is that almost half of the test participants (47%) think that the character's emotional state was happy. For the 14 test participants that watched the Video B first, there were also 2 major groups, 5 that rated the character to be happy and 7 in the "Other" category, where some of the statements were emotions like: "Too happy", "Ready to begin class", "Creepy Happy", "Psychotic, nervous". For the second group the happy emotion can be seen as the 5 that stated the character looked happy or 7 even though 2 test participants wrote creepy or too happy, the basic emotion they both stated, was happy, just a different level of happiness. On a general note for all the answers combined, there were 13 test participants

that rated the happy emotion (41.94%) and 12 that stated “Other” (38.71%), where 2 of these were with the different level of happiness.

Then I had them answer a question about their judgement of the portrayal of the emotion based only on the character’s face.

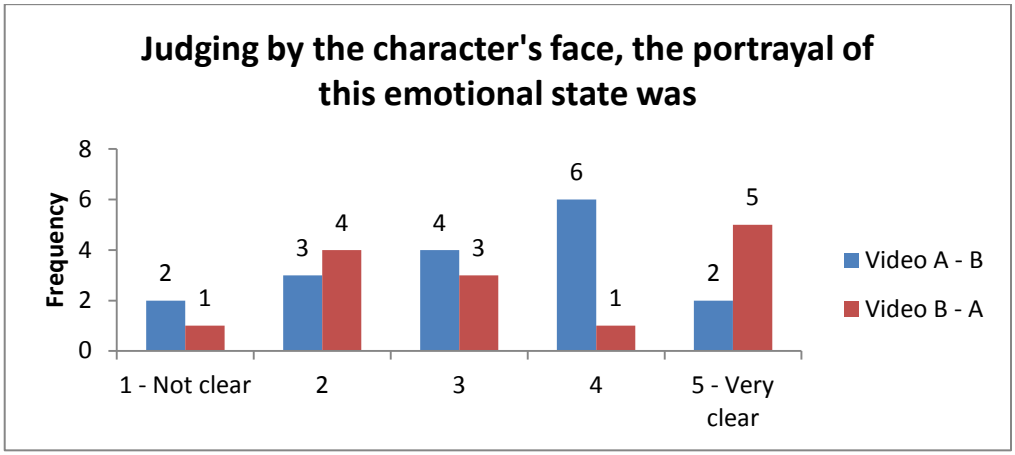


Figure 12 Portrayal of the characters emotional state

As can be seen in figure 12, for the group who watched the Video A first, the answers are a bit more spread out, though still a larger group, 8 test participants (47.06%) that rated it 4 or 5 and agreed that the portrayal of the characters face was clear enough to see his emotional state. For the group who watched Video B, there were 6 participants (42.86%) who gave it a rating between 4 and 5, and 5 (35.71%) who gave it a rating between 2 and 1, so a bit more spread out results compared to the other group. The mean for the first group is 3.18 and the standard deviation for the same group is 1.24, this means that the answers for the first group is centred on the middle, or at least not showing any clear answer of how good the portrayal of the emotional state was. For the other group the mean is 3.36 and a standard deviation of 1.45, and again for this group, seems to be the same for the other group that the answers would lay around the middle. A T-test was also used to see if there were any significance difference between the two groups, and the result showed a p value of 0.36 which is much higher than the 0.05, and therefore there is no significance difference in the two group’s answers for the question about the portrayal of the emotional state.

I asked them to answer a question about how much they agreed with the statement of it was going to be a long day, basically they had to judge the character on his entrance where they could only see the back of the character.

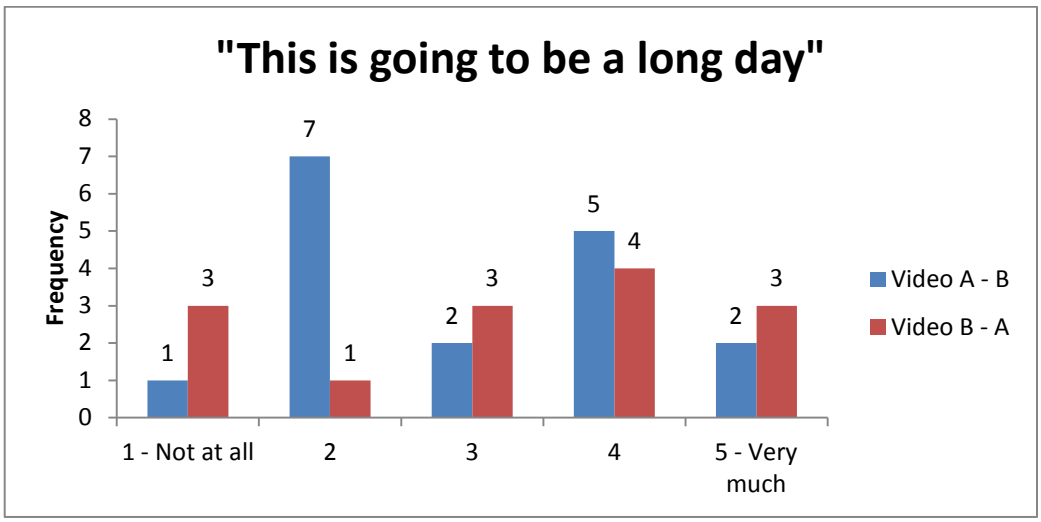


Figure 13 Results of the statement: "This is going to be a long day"

As can be seen in figure 13, for the first group who watched Video A first, that most of them seem to believe that the body language of the character, did not have an impact on their experience of the character, and did not agree with the statement, as 8 (47%) said 2 or 1 and 7 (41%) agreed to with the statement rating it 4 and 5. For the group who watched Video B first, they were more towards agreeing with the statement 7 (50%) rated it 4 or 5, with only 4 (28%) rating it 2 or 1. The mean for the first group is 3 and the standard deviation is 1.22, this means that the average answers would be centred on the middle of the scale with a variation of 1 point each way. For the other group, the mean is 3.21 with a standard deviation of 1.48 and it is again the same case, the average is in the middle of the scale, with a little larger deviation than the other group, but still not showing any clear results. For this question the given p value was found to be 0.33 and again showing there is no significance difference between the two groups.

The next question they were asked to answer was about the effectiveness achieved by the portrayed expression.

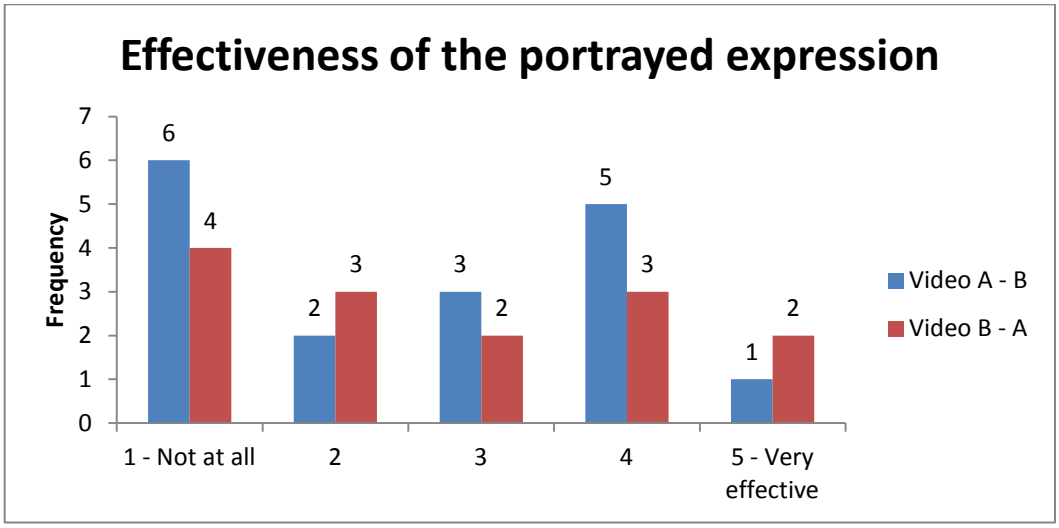


Figure 14 Effectiveness of the portrayed expression

As can be seen on figure 14, from the group who watched Video A, 8 participants (47.06%) rated it 2 or 1, which means that they did not believe it to be very effective, while 6 participants (35.29%) rated it 4 or 5 and thought it to be effective. In the other group, Video B, there were 7 participants (50%) who rated it 2 or 1 and 5 participants (35.71%) that rated it 4 or 5. So a general tendency was to rate it low, as they did not believe that the expression of the character was not effective enough. The mean for the first group is 2.59 with a standard deviation of 1.42, these numbers are closer to indicate an actual result, that in this case the participants tended to rate it low as not being effective. For the other group the mean is 2.71 with a standard deviation of 1.49, and it is the same as the first group, that the tendency was to rate it low. The p value for this question was found to be 0.41 and again showing there was no significance difference in the group's answers.

In connection with the experience of the character I asked them to rate how convincing they thought the expression was.

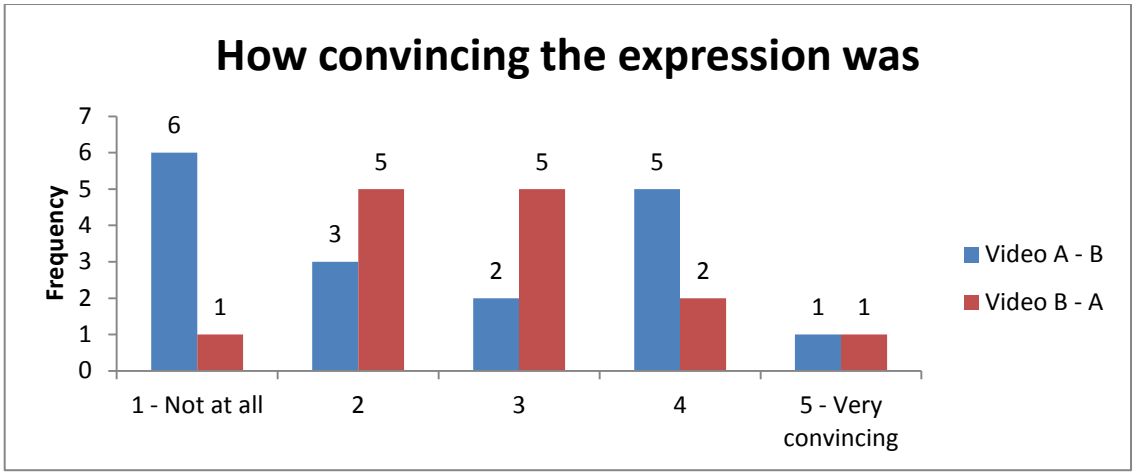


Figure 15 How convincing the expression was

As can be seen on figure 15, the participants who watched Video A first, 9 participants (52.94%) rated it 2 or 1, compared to 6 participants (35.29%) who rated it 4 or 5, and to be convincing. Almost the same happened for the participants who watched Video B, 6 participants (42.86%) rated it 2 or 1, and only 3 participants (21.43%) rated it 4 or 5. Though 5 participants (35.71%) were placed in the middle, saying neither of them and therefore can leave room for speculation if they were going towards being a little convincing or not. The mean for the first group is 2.53 with a standard deviation of 1.42, as with the effectiveness, it seems that the tendency is the same for how convincing the expression was, being rated lower on the scale. For the other group the mean is 2.79 with a standard deviation of 1.05, this group has the clearest result in regards to the standard deviation, as it is close to 1 point difference on the scale, but it can be discussed that the tendency was to rate it lower. For this question the p value was found to be 0.28 and is the closest one to the 0.05 limit, but are still far off and again showing there were no significance difference in the answers for the two groups that participated.

Finally I had them watch each group's second video, and then asked them a single question, which one they believed to be the most realistic, in regards to the facial expression.

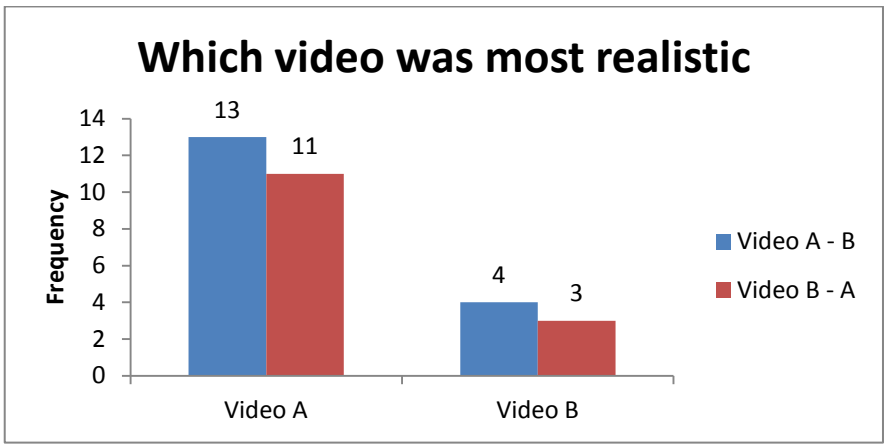


Figure 16 Most realistic video

As can be seen in figure 16, there is 2 big groups from both video groups, which think that Video A looked the most realistic. Of the participants who watched Video A first, there were 13 (76.47%) who said that Video B was the most realistic one. And from the other group 11 participants (78.57%) who also stated that Video A was the most realistic, so for the sake of aiming for realism in the facial expression, the expression library and a little extra work on the upper face, such as the cheeks and around the eyes was sufficient enough to portray a happy emotion.

I also asked them to describe their first reactions to the characters body language and another one to describe their reaction when the character then turned around to face them.

To describe the characters body language, there were some descriptions that were consistent within both groups, that the character seemed to walk slowly and looked sad, somewhat tired and despaired. Even though it was that kind of words used to describe his body language, it still covers some of the same things. I tried to portray the character with a demotivated look, as a demotivated person. You would not walk faster towards something you do not really want to do, and you would most likely hang a bit with your shoulders. There is specially one comment that shows, there is a hint of a link between the student's first look at the teacher and the outcome of their experience, *"Oh no... maybe I should just leave during the first break..."*. By that comment, the body language already had an effect on the viewer, at the moment he saw the teacher coming into the scene, and already there he was thinking about leaving the class to do something else. That also gives an indication of, when

working with virtual agents in applications for learning purposes, that the agents must be welcoming and showing that they “want” to be there and help the students, even though they are not a real person.

Later the test participants should comment on their reaction, when the teacher then turned around to face them.

Here the comments were from happy to scared, angry and some mixed feelings, this could be in relation to the two questions about the effectiveness and how convincing the emotion was. From the comments it could say something about these aspects were not good enough, and therefore there was mixed responses on the reaction to the facial expression. One person also mentioned that the character did not blink with his eyes, and therefore it made it more unnatural for him to look at it. This was something I did not implement as I originally wanted to show the face to the viewer for a short time, so they only got a quick look at his face, and then the video should end. It was decided to give the viewers a little more time to look at his facial expression, and in the process details like these were not implemented, which is of course regrettable that this was not made, as it could have given a more natural look to the character..

Several of the comments, which I received during the test, could relate to the comment that I received on CGSociety, where he took another perspective on the videos, a more artistically viewpoint, and which I believe could have helped me achieve the desired emotion and results.

> ***Hi***
>
> ***I think you’ve got a good idea going here. If you don’t mind, I’ll give you a crit or 3: You should possibly try to exaggerate the body’s animation a little. Also, you should try to add visible brows to the face EG. With black lines. It will improve the emotion a LOT. Since this is a WIP, I’m sure you will improve the rendering and texturing, as the “gray scalish” look it currently has, obscures the details and also hurts the emotion.***

To that comment, I can only agree, after I read it, but as I already was testing and had a lot of test participants already; it is something that I have to use as a future perspective. I think that having more visible eye brows would also help me visualize the emotion, that I tried to portray, and not only using the muscles around the eyes, but also get the visual part with it. But as the character did not come with visible eye brows, I did consider that, when I made the animations. His comment about the textures was a nice critique, as it shows, that it is not only the character as such that is important for having a proper looking environment. It is the general connection between environment and character; it should look like the character belongs in that environment and vice versa.

7.0 Discussion & Conclusion

The motivation to create this project was my general interest in animation, but the character animation in particular. Character animations are very time consuming projects, but the end results of a well done animation can be a very good experience for the viewer. As seen in the animation movies that are being released today. To start off the project, the topic for the initial problem statement was:

How do a virtual agent's facial expressions of emotions affect the user experience in an animated product?

To do this several scientific fields had to be researched, especially in the field emotions. Emotions are a very difficult field as there are many different kinds of facial expressions for the same emotion, as not one person is the same. But still there has been established the so-called six basic emotions, by Paul Ekman. The six basic emotions are Happiness, Sadness, Anger, Fear, Surprise and Disgust. Secondary Emotions were left out of this scope for this project, because the focus was on the visual representation of emotions, and the secondary emotions are more complex, and can be hidden, as a person may feel ashamed as a result of becoming anxious or sad, and this would be difficult to portray.

For the animation part of the project, I believe that manually animating the character's facial expressions could give just as good results, if not better compared to the programmed character behaviour, even though it can take more time to do so. The reason for this is, that the freedom you have when manually animating is much greater than programming the whole sequence. The facial expressions are more personal compared to programming it, as the small details are easier to create manually, and the control of the face is greater.

The animation that was done for this project could also have been better. A few things that lacked, like the work with the eyes and the blinking just shows how complex the work with character animation is, and how many small details that needs to be considered. It is not necessarily a bad thing that these features lacked, as the results showed that the character already had an impact on the experience for some. I believe that with a better animation, I would get more clear results on the experience of the agent.

In conclusion to the results, there was no significant difference in the answers of the test participants, so in that respect there were no clear results of whether the experience of the viewer is changed when looking at the characters facial expression. The majority could recognize that the character was happy, so for the sake of making a facial expression good enough to be recognized was successful. Then for the statistical questions and the actual

test about the effectiveness and how convincing the expression was. Here the results were with a mean lower than 3 and the standard deviation close to 1, which means that the general tendency for these two groups was to rate it not so effective and not so convincing. This could have something to do with the uncanny valley, which I probably ended in, as Keith Sintay describes in a list of, when CGI characters ends up in the uncanny valley (Sintay, 2008);

- Lack of eye jitter (small movements in the eyes when the character looks at something.
- If the eyes always moves with the head, when the head rotates it looks robotic
- -When characters don't blink at all, it looks odd.

The list goes on, but these are the ones that I think have to most impact on my animation, and thus the results, as some commented on the character not blinking at all.

In regards with manually animating or programming the facial expressions he (Sintay, 2008) also stated in his list that:

- Facial features that are mirror-imaged; a lot of people have one side of their face smaller than the other. Features that seem too "perfect".

So this would seem to be another advantage of manually animating facial expressions as there will be natural differences in how you place the different controls in the face, when creating the expressions manually.

The questions that supported the initial problem statement, was the two questions about their first reactions to the characters body language and then the facial expression. Based on these comments, there can be a slight hint that the project as such are working on something that could have an impact on the user experience in an actual application, as one commented on that he was already thinking about going in the first break, after he saw the body language.

If the project would continue, there are several things that need to be fixed, before another test could be conducted, such as setting up the proper environment with more natural looking textures that connects with the character, and then the animation of the character, both the body language and the facial expression in particular needs to be changed.

8.0 Bibliography

A Space Odyssey. 1968. [Film] Directed by Stanley Kubrick. United States of America and United Kingdom: Metro-Goldwyn-Mayer.

Agtivity.com, 2005. *Agtivity.* [Online]
Available at: http://agtivity.com/def/intelligent_virtual_agent.htm
[Accessed 4 April 2012].

Bates, J., 1994. The Role of Emotion in Believable Agents. *Communications of the ACM,* July.pp. 122-125.

Beowulf. 2007. [Film] Directed by Robert Zemeckis. United States of America: Paramount Pictures, Shangri-La Entertainment & ImageMoves.

Beskow, J., 1997. Animation of Talking Agents.

Burgoon, S., 2010. *seanBurgoon.com.* [Online]
Available at: http://seanburgoon.com/goonrig/
[Accessed 2011].

Chatbots.org, 2010. *Chatbots.org.* [Online]
Available at: http://www.chatbots.org/conversational/agent/define_virtual_agent_its_future_objectives/
[Accessed 4 April 2012].

Dautenhahn, K., Bond, A., Cañamero, L. & Edmonds, B., 2002. Creating Relationships with Computers and Robots. In: *Socially Intelligent Agents.* pp. 1-20.

Ekman, P., 1971. Universals and Cultural Differences in Facial Expressions of Emotion. In: J. Cole, ed. *Nebraska Symposium on Motivation.* Lincoln University of Nebraska Press, p. 75.

Ekman, P., 1993. Facial Expression and Emotion. *American Psychologist Association,* 48(4), pp. 376-379.

Ekman, P. & Friesen, W. V., 1971. Constants Across Cultures in the Face and Emotion. *Journal of Personality and Social Psychology,* 17(2), pp. 124-129.

Epic Games, 2006. *Gears of War.* Microsoft Game Studios.

Gallagher, D. F., 2007. *Digital Actors in 'Beowulf' Are Just Uncanny.* [Online]
Available at: http://bits.blogs.nytimes.com/2007/11/14/digital-actors-in-beowulf-are-just-uncanny/
[Accessed 14 May 2012].

Ghost VFX, 2009. *Tulip | Bear.* [Online]
Available at: http://ghost.dk/?p=148
[Accessed 26 March 2012].

I, Robot. 2004. [Film] Directed by Alex Proyas.

Jdoubles, 2007. *CreativeCrash.* [Online]
Available at: http://www.creativecrash.com/maya/downloads/character-rigs/c/the-andy-rig

Martin, J.-C. e. a., 2011. Coordinating the Generation of Signs in Multiple Modalities in an Affective Agent. In: Paris. p. 19.

Peters, C., 2011. Editorial: "Emotion in Interaction". In: P. Petta, ed. *Emotion-Oriented Systems.* Coventry. sp. 5.

Picard, R. W., 1995. *Affective Computing,* Cambridge: M.I.T Media Laboratory.

Rio. 2011. [Film] Directed by Carlos Saldanha. United States of America: Blue Sky Studios and Twentieth Century Fox Animation.

Sintay, K., 2008. *Animation Mentor.* [Online]
Available at: http://www.animationmentor.com/tiptrickblog/author-KeithSintay.html
[Accessed 14 May 2012].

Steele, K. et al., 2009. Markerless Versus Marker-Based Motion Capture: A Comparison of Measured Joint Centers. August.

Studios, B. G., 2011. *The Elder Scrolls: Skyrim.* Bethesda Softworks.

Techtarget.com, 2003. *Techtarget Definition of Virtual Agent.* [Online]
Available at: http://searchcrm.techtarget.com/definition/virtual-agent
[Accessed 4 April 2012].

Techtarget.com, 2005. *Definition: Affective Computing.* [Online]
Available at: http://whatis.techtarget.com/definition/0,,sid9_gci922555,00.html

Terzopoulos, D. & Mones-Hattal, B., 1997. Facial Animation: Past, Present and Future.

Thomas, F. & Johnston, O., 1981. *The Illusion of Life.* New York: Walt Disney Productions.

Tinwell, A., Grimshaw, M., Nabi, D. A. & Williams, A., 2011. Facial Expression of Emotion and Perception of the Uncanny Valley in Virtual Characters. *Computers in Human Behavior,* March, 27(2), pp. 741-749.

Toy Story. 1995. [Film] Directed by John Lasseter. United States of America: Pixar Animation Studios and Walt Disney Pictures.

Tull, M., 2009. *Secondary Emotions.* [Online]
Available at: http://ptsd.about.com/od/glossary/g/secemotdef.htm
[Accessed 12 April 2012].

Williams, R., 2001. *The Animator's Survival Kit.* Queen Square London: Faber and Faber Inc.

Yan, J. & Agada, R., 2010. Life-Like Animated Virtual Pedagogical Agent Enhanced Learning. *Journal of Next Generation Information Technology,* 31 August.

APPENDIX A

This is the results that were gathered from the first iteration of the testing phase, the 10 participants were asked to describe the characters facial expression.

First Iteration										
Based on the characters facial expression, the character looks:										
Happy		Sad		Indifferent		Other:				
6		0		1		Clueless				
						Creepy / slightly insane				
						Eyes looking slightly sad / mouth contempt				

APPENDIX B

For the second iteration, there were problems with the quality of the videos, so the test was stopped and then went back to fix the quality of the videos, but also some of the comments that they had about the character. First Excel spread sheet is from the Video 1.

Timestamp	Looking at his face, how would you describe the character's emotional state?	Judging by the character's face, the portrayal of this emotional state was	If you were the student and the character was the teacher, how much would you agree with the following statement	If the purpose was to portray a happy emotion, how effective was the character's expression?	If the purpose was to portray a happy emotion, how convincing was the character's expression?	When the teacher entered the scene, how was your initial reaction to his body language?	What was your reaction when he turned around, facing you?	Which video would you consider to look the most realistic?
4-29-2012 12:39:08	Indifferent	2	5	1	4	mistænksom	han ved noget	Video 1
4-30-2012 14:52:34	scouting the room for something/someone	4	5	1	1	that he is another old, simpathetitc teacher. grandpa like	uhm: "oh, crap!"	Video 1
4-30-2012 15:17:21	kinda surprised I'd say, kinda hard to see because of the quality	1	2	2	1	he looked kinda bored or like he was tired and just wanted to get it over with ^_^'	I lol'ed, due to the difference between what you expected from his posture and his face expression. but I must say that I have no idea what is the purpose here, I can't really tell if he's surprised, sad, happy, laughing or anything else.	Video 1
4-30-2012 16:23:30	Indifferent	1	1	1	1	sluggish, slow	the video quality was to hard to tell	Video 1
4-30-2012 16:24:00	Can't really see it - it's too small...	1	3	1	1	His entrance seemed very slow, deliberate	I really could not decipher his facial expression whatsoever	Video 1
5-1-2012 9:26:43	paranoid	1	3	1	1	depressed, slow, hessitent and stiff	expecting he would tell that the world would end or something like that, nothing good with that facial expression *the next question none of them is realistic, one(#1) is paranoid/scared and the other(#2) is overly enthusiastic (and that is bordering psychotic)	Video 1

Next spread sheet is from Video 2 and their answers to the questions

Timestamp	Looking at his face, how would you describe the character's emotional state?	Judging by the character's face, the portrayal of this emotional state was	If you were the student and the character was the teacher, how much would you agree with the following statement	If the purpose was to portray a happy emotion, how effective was the character's expression?	If the purpose was to portray a happy emotion, how convincing was the character's expression?	When the teacher entered the scene, how was your initial reaction to his body language?	What was your reaction when he turned around, facing you?	Which video would you consider to look the most realistic?
4-29-2012 12:59:06	Despairing - so close to sad	2	4	2	2	Slow movement so seem kind of sad	Although he smiled, his eyes combined with his posture made him seem more sad and despairing.	Video 2
4-30-2012 14:27:58	Happy	3	3	4	3	in some way he seemed frightned	non	Video 2
4-30-2012 15:35:39	Crazy	4	3	4	4	I could read anything from it.	That he had a crazy smile.	Video 2
4-30-2012 16:46:37	He is smiling (happy) but the eyes are making him look creepy.	1	3	2	3	Normal, maybe looking a bit tired.	He was leaning over the table looking angry. But his eyes where making him look really creepy.	Video 2
4-30-2012 17:48:29	Sad	2	3	2	2	Slow = not that confident, kind of sad	I was surprised that he was smiling, and thought he looked kind of creepy, since his eyes look sad but he is smiling, and the way he looks around, just smiling when seeing the students	Video 2
5-1-2012 19:04:43	A little crazy - looks like he was "pushed over the edge", or had some unsettling news right before class, so he's trying to fake-smile his way through it.	1	3	2	3	He walked slightly hunched, so I figured he was tired - didn't want to be there.	Fake-smile! His eyes seem to indicate he's sad or unsettled, so I kind of felt bad for him, having to lecture when he clearly didn't want to.	Video 2
5-3-2012 12:52:08	Insane	4	2	3	2	Quite neutral..	He looked insane, staring across the room with an unrealistically happy smile. P.S. You made a small flaw in the naming of the videos.. The first video was called "Video 2" on youtube while the second was called "Video 1".. I assume people in general would say that the answer "Video 2" is actually the one called "Video 1" on youtube.. I have answered as such.	Video 2
5-10-2012 12:35:25	Happy	2	4	2	2	Sad, boring, tired	suprising because i expected a sad face because of his body language	Video 2

APPENDIX C

For the final test, I had 31 test participants, and their answers look like this. For Video A:

Timestamp	Looking at his face, how would you describe the character's emotional state?	Judging by the character's face, the portrayal of this emotional state was	If you were the student and the character was the teacher, how much would you agree with the following statement	If the purpose was to portray a happy emotion, how effective was the character's expression?	If the purpose was to portray a happy emotion, how convincing was the character's expression?	When the teacher entered the scene, how was your initial reaction to his body language?	What was your reaction when he turned around, facing you?	Which video would you consider to look the most realistic?
5-9-2012 16:50:56	Happy	5	2	4	4	He looked very relaxed	There is a happy banana :)	Video A
5-9-2012 20:05:16	Sad	3	4	1	1	he seemed to walk fairly slow	I tried to understand his facial expression.	Video B
5-10-2012 10:55:35	Happy	4	2	3	3	He had his back turned.	A bit intimidating when he leaned over the table.	Video A
5-10-2012 11:06:56	Happy	4	5	4	4	Moving a bit slow, very skinny	Scared, frigthened, intimidated	Video A
5-10-2012 12:09:35	Awkward	4	2	1	1	Not good.	Confused, like someone told a joke I didn't get.	Video A
5-10-2012 12:15:05	Happy	4	2	4	2	Felt he/she was meticulous and robotic, stern and emotionless.	Almost scheming, not really happy, but obvious that was the emotion attempted.	Video A
5-10-2012 13:09:56	up to something	3	2	2	2	His head in the clouds, thinking.	"This can't be good"	Video A
5-10-2012 14:04:38	insane	1	3	2	2	slouched, slow moving and heavy, I regarded him as sad	confusion. smiling mouth with lifeless eyes and brow (wide-eyed with no percievable change of shape, the pupil movement was also quite alien) didnt clearly convey any emotion.	Video A
5-10-2012 16:37:24	Angry	3	4	1	1	he looked tired/non-caring	he looked angry	Video A
5-10-2012 16:40:45	Happy	4	2	4	4	Very relaxed and laid back	That he seemed to be in a good mood	Video A
5-10-2012 17:08:52	nærmest lidt nervøs	2	4	3	3	han virkede rimelig laid back	lige først lignte det et smil men det virkede mere som en nervøsitet efter lidt tid	Video A
5-10-2012 17:30:13	Happy	2	3	4	4	Han virker lidt opgivende ?	Var nødt til at big screene den , før jeg kunne se at han smilte. Min første reaktion var at han, igen, virkede opgivende, og en smule aggressiv ..	Video A
5-10-2012 18:03:08	Angry	2	5	1	1	He was relaxed	Confusion	Video B
5-10-2012 19:58:09	Happy	4	2	3	4	Posture seemed sad - shoulders down, hunched back and looking down	of cause the video loaded the same moment, but it was a surprise that he turned and faced the class, hands on the desk, as if he wanted to get the attention of the class	Video A

5-10-2012 20:19:20	Angry	1	4	1	1	Normal	Seemed angry	Video A
5-13-2012 11:42:48	Unstable	3	4	1	1	The way he kind of sulked looked sad	The body movement seemed somewhat sad, but his face was just plain creepy. A very unnerving and unnatural smile combined with the unblinking eyes of Sauron and long sweeping movements of his head.	Video B
5-13-2012 19:09:02	Happy	5	1	5	5	Sad	smiling and happy	Video B

And for Video B the test participants answered.

Timestamp	Looking at his face, how would you describe the character's emotional state?	Judging by the character's face, the portrayal of this emotional state was	If you were the student and the character was the teacher, how much would you agree with the following statement	If the purpose was to portray a happy emotion, how effective was the character's expression?	If the purpose was to portray a happy emotion, how convincing was the character's expression?	When the teacher entered the scene, how was your initial reaction to his body language?	What was your reaction when he turned around, facing you?	Which video would you consider to look the most realistic?
5-10-2012 10:51:43	Happy	5	1	5	5	Walked like an old man	Positive....funny	Video A
5-10-2012 10:59:59	Ready to begin class	4	1	3	3	Attentive	Ready to hear what he is going to say	Video A
5-10-2012 11:05:56	Angry	5	1	1	1	I thought he looked depressed.	I did not react very much to it. The figure looked mad.	Video A
5-10-2012 11:43:37	too happy	5	5	4	4	neutral	serial killer	Video A
5-10-2012 13:53:30	Happy	5	3	4	4	Not the best posture walking in. Standing by the table the body language was welcoming.	I was slightly curious to see what would happen next. Comment: "Which video would you consider to look the most realistic?" this question makes no sense. The difference in facial expression does not make me feel that one video is more realistic than the other. If I could have chosen "equally realistic" or "equally unrealistic" I would have. I choose B simply because I have to choose one or the other. Good luck with the project :)	Video B
5-10-2012 13:54:14	Overexcited	5	4	2	2	He looked a bit slouchy, I got the feeling he's meant to be some old professor...	...but then he turned around with this Joker smile on his face and I was like 'this guy has either won a few millions or is just planning to go on a killing spree'	Video A
5-10-2012 16:24:03	Deranged madman	3	4	2	2	He falling asleep?	AAAAAHHHHHH!!!! ! The joker from batman!!!	Video A
5-10-2012 16:51:36	Happy	3	3	5	3	going to desk - he looked like he was sad. after turing - he is trying to convince someone of something	i don't like him. he's trying to convince someone of something	Video A

5-10-2012 17:32:54	Happy	2	3	4	3	Hans holdning og kropssprog virker til at han er meget bestemt og ikke lystig og glad som hans ansigt ligger op til.	Jeg blev ikke mere overbevist om at han var glad, selvom han smilte fra øre til øre.	Video A
5-10-2012 21:41:44	Creepy happy	2	4	2	2	He was leaning forward with this very creepy big smile looking around the class room, made me feel like he had some evil plan/sceme to carry out...	The big smile looking around the class room, made me feel like he had some evil plan/sceme to carry out...	Video A
5-11-2012 10:16:48	mystory	2	5	1	2	Stupid, good for nothing teacher.	Mysterious, evil and fake laugh that might give a sign for bad exam results etc.	Video B
5-11-2012 13:53:29	Sad	3	4	1	3	Oh no... maybe I should just leave during the first break...	Spooked.	Video A
5-13-2012 9:58:52	Happy	1	5	3	2	It was ok for a teacher	I got scared by his face expression because of the eye gazing	Video B
5-13-2012 11:10:25	Psychotic, Nervous	2	2	1	3	Stiff, not enough weight in his steps	"Here's Johny"	Video A