



Universiteit Utrecht

Is FaceReader an Applicable Tool to Measure Facial Expressions? A comparison between FaceReader and Facial EMG.

Clinical and Health Psychology Master's Thesis

Supervisor: Marloes Eidhof, MSc

Selen Urgancioglu

5951860

Utrecht University

July 2017

ABSTRACT

The present study researched if the facial expressions could be measured effectively by FaceReader Noldus software by comparing FaceReader and facial EMG. The study consists of participants watching 4 different movie clips evoking 3 different emotions; sadness, disgust, fear and one neutral movie clip. Participants watched each movie clip for 5 times in a randomized order while they had measurement instruments for EMG on their faces and the camera was recording their faces for FaceReader. They answered VAS questions at the end of the experiment which included ranking the emotions they felt on a scale from 0 to 10, during the first time and the fifth time they watched each movie clip. VAS gave us the opportunity to analyze the emotions the participants thought they felt, apart from what had been measured from their facial expressions and if the outcome was consistent with FaceReader and EMG. Also, a comparison had been done between the facial expressions of participants during the first time and the last time they watched the movie clips to analyze the habituation effect. The outcome suggested that there was a significant difference between the FaceReader and EMG's outcome. In contrast to VAS outcome, FaceReader's measurement suggested that there was not a significant difference between the first time and the fifth time the participants watched the videos. Further improvements and researches are needed before being able to state that FaceReader could be used in clinical practice.

Table of Contents

| | |
|--|----|
| ABSTRACT..... | 2 |
| INTRODUCTION..... | 4 |
| 1. THEORETICAL BACKGROUND | |
| 2. RESEARCH QUESTIONS AND HYPOTHESIS OF THE STUDY | |
| METHOD..... | 8 |
| 1. SAMPLE | |
| 2. MATERIALS | |
| 3. DESIGN | |
| 4. PROCEDURE | |
| 5. PLAN OF ANALYSIS | |
| RESULTS..... | 11 |
| 1. FACEREADER COMPARED TO EMG | |
| 2. FIRST MEASUREMENT COMPARED TO FIFTH MEASUREMENT | |
| 2.1 FACEREADER | |
| 2.2 EMG | |
| 2.3 VAS | |
| DISCUSSION..... | 16 |
| BIBLIOGRAPHY..... | 22 |
| APPENDIX..... | 25 |

INTRODUCTION

After exposure of a traumatic event such as; death, threatened death, actual or threatened serious injury, or actual or threatened sexual violence, people might develop physical or psychological problems. According to DSM-5, these problems might include disorders like; reactive attachment disorder, disinhibited social engagement disorder, posttraumatic stress disorder (PTSD), acute stress disorder, and adjustment disorders. About 10% of people might develop post-traumatic stress disorder (PTSD) after experiencing a traumatic event. (Stroebe & Schut, 2001)

Some individuals with PTSD will seek for treatment and over the centuries many different therapies have been developed in order to decrease or help people cope with their symptoms. Sometimes these individuals report problems with expressing or even understanding their own emotions. It is very common for individuals with PTSD to think that they are less able to feel positive or negative emotions and they might also have problems with emotion regulation. (Ehring & Quack, 2010) Also, some people do feel emotions but have difficulties with expressing the emotions to others. They might have difficulties with telling how they feel. (Nightingale, & Williams, 2000) In these cases it is difficult for the patient and therapist to get a clear picture of what the patient is feeling. So, understanding one's own and each other's emotions is very important, but we are not always capable of doing this. This is the topic of this thesis: Are there alternative ways to measure the emotions without having to put them into words?

“The young and the old of widely different races, both with man and animals, express the same state of mind by the same movements.” (Darwin, 1872). Since Darwin proposed that the facial expressions have a biological basis and that they are universal in his book *The Expression of the Emotions in Man and Animals* (1872), the idea that facial expressions can be scientifically measured started to develop. Facial expressions are the key to understand the emotions when the words are not telling what is going on inside of the mind. *Paul Ekman* (1992) emphasized the importance of facial expressions. He showed that individuals with congenital facial paralysis, who are unable to show any facial expression, are having a lot of difficulties in developing interpersonal relationships.

If facial expressions are a valid representation of emotions, this would enable us to measure emotions with measuring facial expressions. Yet, an important question is: “Are facial expressions a valid representation of emotions? ?” and “Are the facial expressions universal or cultural?”

Whether or not facial expressions are universal was studied by Darwin (1872), but it became a popular research question after late 1960's (Ekman, 1993). The several studies were conducted by Ekman(1972), Ekman & Friesen (1971), Izard, (1971), and Matsumoto (2011) in this research field. Charlesworth (1973), found in his experimental study that the facial expressions of children that were blind, were similar for children that were not blind. This is evidence for who are blind and who can see have similar emotions and similar ways to express them. According to this information, we can assume that observation and environmental factors are not very effective on facial expressions and that they might be universal. Per the meta-analysis of Matsumoto, Keltner, Shiota, Frank and O'Sullivan in 2008, there are 75 studies done by different researchers around the world proving that the same facial expressions occur when the emotions were triggered spontaneously. As a response to the question we mentioned, this was counted as strong evidence to determine the universal facial expressions for the 7 emotions; happy, disgust, anger, fear, joy, contempt, sad. (Matsumoto, 2011)

The research questions of this thesis are mainly based on how could emotions and facial expressions could be measured and alternative methods that could be used. The next questions are “How can we measure facial expressions? and “What scientifically proven by empirical evidence methods are already used to measure the emotions?”

There are various methods developed over time to measure emotions and each of them have different benefits and disadvantages. The first one is the Facial Action Coding System, developed by Ekman (Ekman& Friesen, 1978) was published as a manual and then it was developed and published again in 2002. (Wolf, 2015) Facial Action Coding System is an anatomical based guide which allows you to make a description and categorization of the facial expressions according to the movements of the facial muscles. (Wolf, 2015)

Electromyography which is one of the main focuses of this thesis is a method that is used to measure the movements of muscles and the nerve cells too. There are electrodes used as measurement instruments of the EMG and they connect the muscle and the EMG. Then once the EMG is connected to the electrode, it measures the activity of the muscle area it is connected to by detecting the electrical signals. Then the EMG shows this collected data to the researcher as graphics and as numerical data. (Wolf, 2015). Electromyography can be done with intramuscular electrodes or surface electrodes. (Turker, 1993)

Facial EMG is a variety of electromyography that is used to measure and differentiate the emotional reactions given by the people per the movements in their facial muscles because of

facial expressions. It is applied through surface electrodes. There are a lot of studies used facial EMG meanwhile the participants are looking at pictures or movies or listening to music pieces. Those studies showed that facial EMG captures the facial expressions and mood changes. (Tan, Jun-Wen et al, 2011)

Electromyography is a method used by numerous researchers and most of the researchers agree on that it is reliable, valid, stable and the accuracy of the results is empirically proven. (Wolf, 2011)

Yet facial EMG is an effective method to measure facial expressions, it has some challenges indeed. The EMG machine and all the necessary devices such as electrodes are expensive. Also, using EMG requires a clinician who knows and is experienced in using the EMG and doing the analyses afterwards. There is also the preparation process which includes explaining the process to the client, preparing the electrodes and placing them correctly and then spending time on analyzing them. EMG is not available to be used anywhere, it needs an environment where all the devices can be placed like a laboratory. So, to conclude we might say that EMG is an accurate and sensitive method to measure the facial expressions through muscles but it also consumes the time, energy, and money of the clinician.

In consideration of the challenges about facial EMG, I want to mention one last method to measure facial expressions, which constitutes the base of this thesis, FaceReader.

FaceReader is a software, developed for facial expression analysis. FaceReader can analyze the universal emotions determined by Ekman such as; happy, sad, scared, disgusted, surprised, angry, and it can also recognize neutral and contempt expressions. (Ekman, 1992) In fact, FaceReader has an extensive usage area such as; psychology, development, education, marketing, consumer behavior may make use of the software. However, in this thesis we will be focusing on the usage of FaceReader in clinical psychology.

FaceReader's working principle is; first it recognizes the face, then it contributes a 3D model and does the analysis on the model and gives the results on a scale from 0 to 1. (Lewinski, 2014). To get the face recognition for the FaceReader, a webcam or an eye tracker is used. Also, FaceReader online can be used through people's own webcams and it allows to get the visual data without the restriction of time and location. (Lewinski, 2014) This is a great advantage for FaceReader especially compared to EMG, as it is not expensive, and it could be used everywhere with a computer and a web camera.

There question we aim to answer is: “Is FaceReader an applicable tool to measure emotions?” It is not yet empirically proven that FaceReader might be an instrument that can be used by clinicians to measure facial expressions. If this study reveals that FaceReader can accurately measure emotions, it could be tested in clinical practice and it would be a cheaper, faster, and easier method than EMG.

In this thesis, we did a comparison between the two measurement methods mentioned; Facial Emg and the FaceReader. The reason we compared those two methods is that the efficacy of EMG is scientifically proven but the FaceReader is not, at least until now. By comparing their results, we could see if their results are coherent with each other and if the FaceReader is reliable. Our hypothesis was “Facial expressions of an individual for the emotions; disgust, fear, sadness can be measured accurately by FaceReader and the results would be coherent compared to facial electromyography.” Also, a comparison of the facial expressions of the participants during the first and fifth time they watched the videos was done to see the habituation effect, the second hypothesis is; “There is a significant difference between the facial reactions of people between the first time and the fifth time that they watch a movie clip which is triggering their emotions.”

METHOD

Sample

Flyers were prepared to announce the experiment with the contact information and how to sign up for it. We put those flyers to different places like library, cafeteria in Utrecht University to find participants. Also, a poster was prepared and shared on social media with the same information.

Our total sample consisted 33 participants and our final sample was 26 participants because we had to exclude 7 participants whose outcome results did not satisfy our several criterias. Some participants did get out of the angle of the camera during watching the videos and this enabled the FaceReader to measure. Besides, some participants' EMG results were not valid, perhaps caused by a problem in stiffness of the electrodes and we had to exclude them as well.

76% of the final 26 participants of the study were Dutch, 9% of them were German, 3% of them were Spanish, 3% of them were Irish, 3% of them were Vietnamese, 3% of them were Norwegian and 3% of them were British. The age of the participants changes between 18 and 30.

The sample of this experiment only consisted of women. Only women participants were included in the study because it is known by scientific evidence that men exhibit less emotion and facial expressions than women. (Montagne, Barbara, et al, 2005) In this study, there were only women participants who were not wearing glasses or who can see the screen well when they take them off. The reason for this restriction was because the glasses were covering an important part of the face and this fact might affect the angle of view of the web camera we were using for FaceReader and the analysis that will be done by FaceReader.

Materials

Informed consent forms were used with detailed information about the study and the procedure. (See Appendix A) Computers were used for participants to watch the videos and to control the experiment. Moreover, a demographic questionnaire was used to get information about the participants' backgrounds like; their age, nationality, language, medical condition, their familiarity with EMG and videos before they participated to the experiment and about their comments. (See Appendix B)

Movie Clips

4 different short video clips were determined triggering different emotions as; fear, sadness, disgust and neutral for participants to watch. The neutral video was added to the experiment as a control mechanism. The videos we used were videos that have been proven to trigger that specific emotion in previous researches. The fear clip was a scene from the movie "Copycat" and the neutral clip was a low emotion arousing scene from the movie "The Lover". (Schaefer, Nils, Sanchez, & Philippot, 2010) The disgust clip was a scene from the movie "Pink Flamingos" and the sadness clip was a scene from the movie "The Champ". (Gross, & Levenson, 1995)

FaceReader

Facial expressions were recorded using the Noldus FaceReader software. For recording, the camera Sony SRG300H 30x 1080p/60 HD PTZ Camera was used.

EMG

EMG required its special measurement instruments as electrodes. Facial EMG activity was recorded bipolarly using sintered Ag/AgCl electrodes. The diameter was 4/8/12 mm. EMG

signals were logged with MindWare software (EMG 3.0.2.1; MindWare Technologies, Gahanna, OH). Notch Filter was 50 Hz, Sampled at 1000 Hz, Band Pass was 20 Hz-200Hz. The rectifying and smoothing of the signals were done with a 20 Hz low-pass filter with a time interval of 100 ms.

VAS

There is a questions part after the videos are finished, which is referred as VAS questions, the participants are supposed to rank on the screen how much watching the first time and the last time each video clip triggered their emotions.

Design

Four different video clips, one for triggering disgust, one for sadness, one for fear and one neutral video clip which does not activate any emotion particularly were watched by the participants. The videos last approximately 30 seconds. Each video clip was showed for 5 times, to see the habituation effect, but they were not shown in a row. The video clips were shown in a randomized order for each participant. The experiment takes about 20-25 minutes in total including the preparation process.

Participants are supposed to watch the video clips meanwhile they have electrodes on their face for EMG and with the recording of the camera on top of the computer for FaceReader. Thus, their reactions could be measured simultaneously by the two methods we wish to compare. At the end of the videos, participants see questions on the screen and they answer the VAS questions. Later, the participant is supposed to answer the demographic questionnaire.

After analyzing the data, we obtained, we will be able to compare the results of FaceReader, EMG and Vas scores and validate or reject the hypothesis.

Procedure

The first thing the researcher did when we met the participant at the lab is to tell about the procedure. The participants are told that they will be watching short video clips meanwhile measurement instruments for the EMG would be placed on their faces. They were asked to read the informed consent and then they sign it. Following, the researcher prepared the measurement instruments and placed them on the participant's face. To prepare the electrodes, the special stickers should be pasted on the electrodes and they should be renewed for each participant. Then, the special gel should be applied on the sticker to be able to transmit the electrical signals.

The electrodes were placed on the facial area that we want to measure the activity from. Since we are interested in the facial expressions triggered by the emotions; fear, disgust, and sadness, we placed our electrodes to the facial areas which are activated during those emotions are felt according to the EMG guidelines. We had two electrodes for each emotion, so 6 in total. Two electrodes were placed in forehead; above the start of left eyebrow, to measure fear. Two electrodes were placed next to the left nostril, to measure disgust. Two electrodes were placed under the left corner of lower lip to measure sadness.

After sticking the electrodes and controlling if they are stable, the participant was guided to the lab and the researcher informed the participant that he will be present in the lab as an observer during the experiment. Later, the researcher connected the electrodes to the EMG and explained the VAS questions and that the instructions will appear on the screen at the end of the videos. It was followed by explaining the fact that participant and the observer must not talk during the experiment, but the participant could stop the experiment anytime she wants, the researcher started the recording of the FaceReader and this started the experiment. The observer sat on a chair which is in the blind spot of the camera's angle. The participants watched the videos for 13 minutes and then answered the VAS questions from the computer. At the end of the experiment, the recording for the FaceReader was stopped and the electrodes were removed by the researcher.

Finally, the participant was supposed to answer the demographic questionnaire given by the researcher. Besides the questions about their background, it was also asked if they had facial paralysis or a similar condition, if they were using any medication that could interfere with their facial reactions or emotions, if they had botox or plastic surgery and if they had visited the dentist and they used any anaesthetics there in the last 48 hours. These questions were asked because those medical conditions might affect the ability to show facial expressions and this might affect the outcome. These questions were all very important to reduce the external factors and for the reliability of the results. The purpose of the study was explained to the participant at the end of the experiment.

Plan of Analysis

We examined the effectiveness of FaceReader in analyzing the facial reactions, by comparing whether their outcomes were coherent with EMG in the first-time points. We also

compared the outcome of FaceReader, EMG and VAS scores for the first and fifth time points of measurement to understand if the participants' expressions changed between the first and fifth show of each video.

Paired sample t-tests were conducted to compare FaceReader's and EMG's measurement on first time points for sadness emotion in sad video, disgust in disgust video and fear in fear video. Paired sample t-tests were also conducted separately to compare FaceReader's, EMG's, and VAS's first time and fifth time measurements for sadness, disgust and fear.

We conducted Helmert contrasts for FaceReader, EMG and VAS.

RESULTS

Paired Samples t-tests

We conducted paired sample t-tests to compare the outcome of FaceReader and EMG to be able to validate or reject the hypothesis. We also conducted paired sample t-tests to compare the outcome of two different time points of measurement.

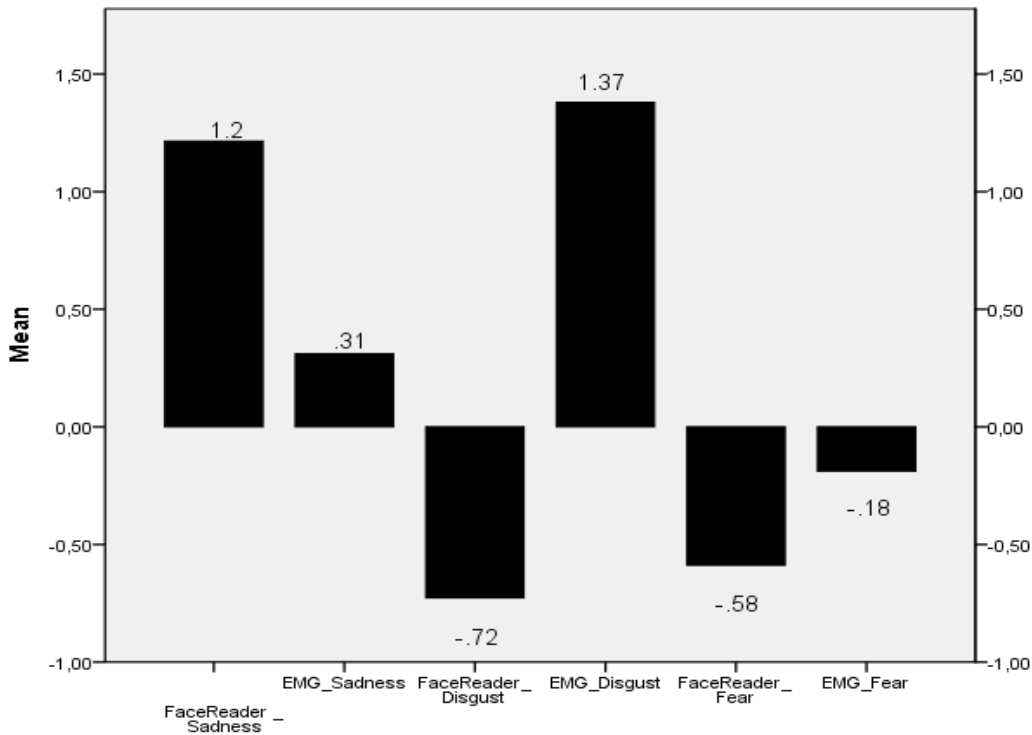
Face Reader vs EMG

Two-tailed paired samples t- tests were conducted to compare the measurement of FaceReader and the measurement of EMG for the sadness emotion in the sad video, for disgust in disgust video and for fear in fear video for the first-time point. For sadness, there was a significant difference in the scores for FaceReader (M= 1.2, SD= 0.59) and EMG (M= 0.31, SD= 1.28) conditions $t(25) = 3.49$ $p=0.002$ $p<.05$.

For disgust, there was a significant difference in the scores for FaceReader (M= 0.72, SD= 0.17) and EMG (M= 1.37, SD=1.99) conditions $t(25) = 5.50$ $p= 0.000$ $p<.05$.

For fear, there was a significant difference in the scores for FaceReader (M= 0.58, SD=0.13) and EMG (M= 0.18, SD= 0.82) conditions $t(25) = 2.46$ $p= 0.021$ $p<.05$.

FACEREADER vs EMG

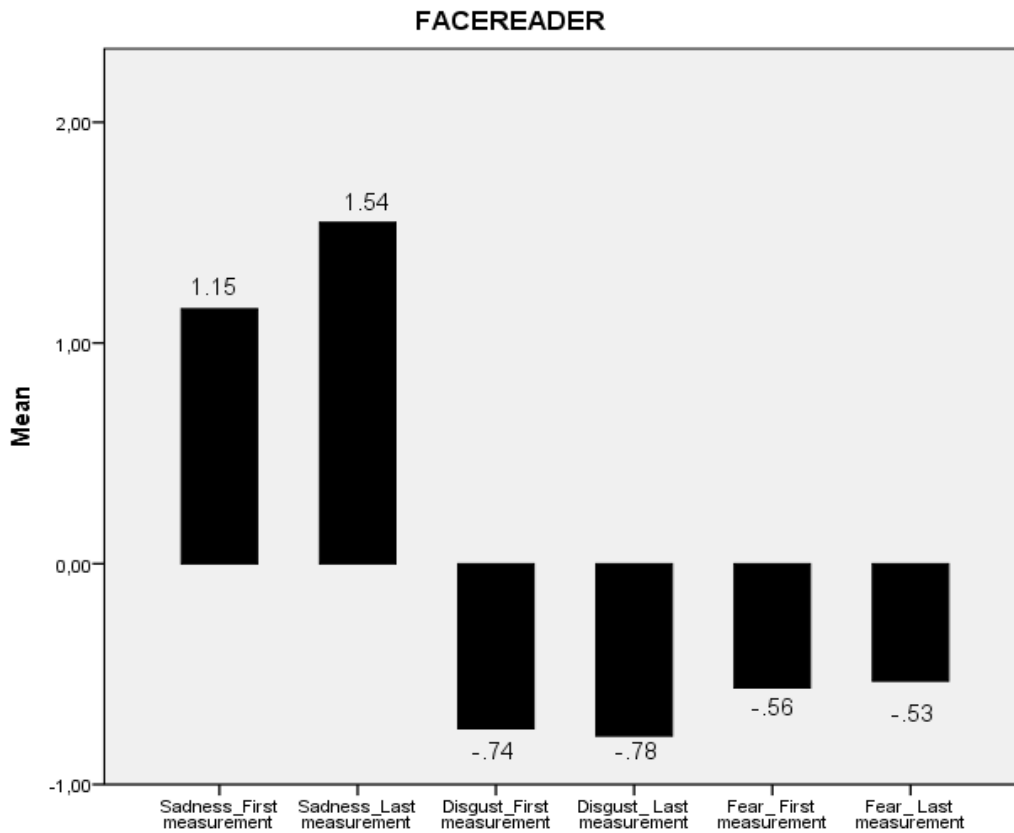


FaceReader- First Measurement vs Last Measurement

Two-tailed paired samples t- tests were conducted to compare the measurement of FaceReader in the first-time point and in the last time point for the sadness emotion in the sad video, for disgust in the disgust video and for fear in the fear video. In the sadness video, there was not a significant difference in the scores for the first time (M= 1.15, SD=0.603) and for the last time (M=1.54, SD=0.604) conditions $t(28) = 2.03$ $p = 0.052$ $p > .05$.

For disgust measures, there was not a significant difference in the scores for the first time (M= 0.74, SD=0.18) and for the last time (M=0.78, SD=0.13) conditions $t(28) = 1.45$ $p = 0.158$ $p > .05$.

For fear in the fear video, there was not a significant difference in the scores for the first time (M= 0.56, SD=0.15) and for the last time (M=0.53, SD=0.21) conditions $t(28) = 0.88$ $p = 0.382$ $p > .05$.

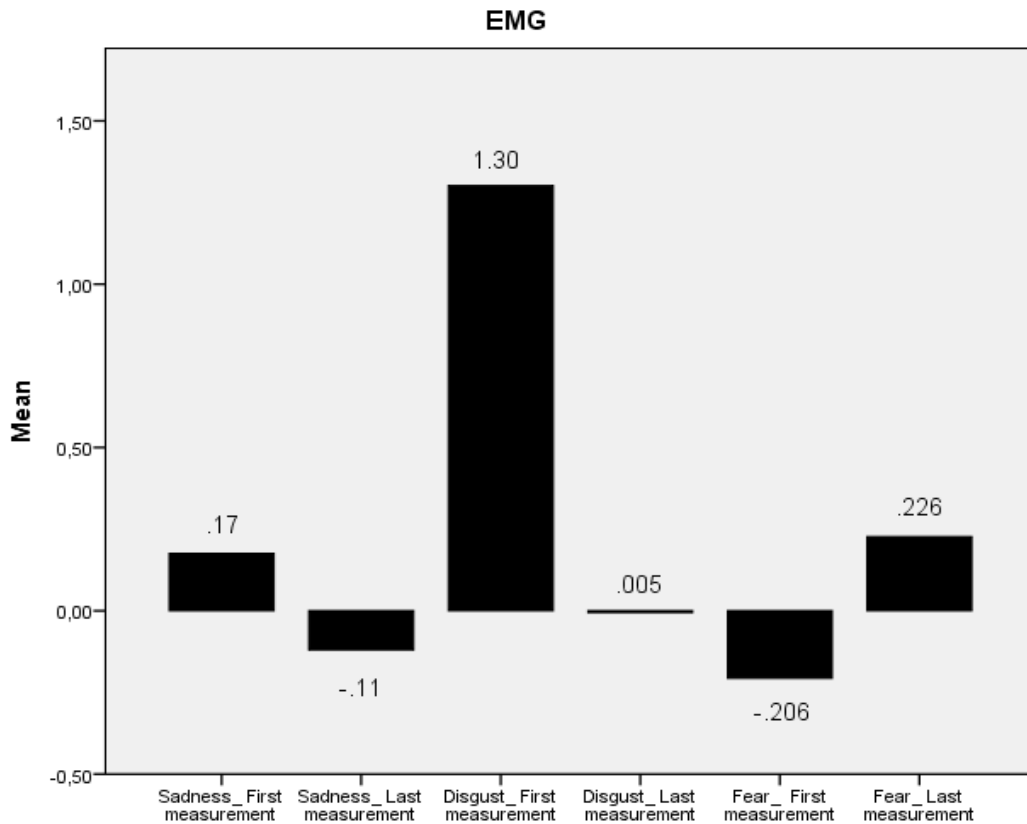


EMG- First Measurement vs Last Measurement

Two-tailed paired samples t- tests were conducted to compare the measurement of EMG in the first-time point and in the last time point for the sadness emotion in the sad evoking video for disgust in the disgust video and for fear in the fear video. For sadness in the sad video, there was not a significant difference in the scores for the first time (M= 0.17, SD= 1.25) and for the last time (M= 0.11, SD= 0.79) conditions $t(29) = 1.17$ $p = 0.250$ $p > .05$.

For disgust in the disgust video, there was a significant difference in the scores for the first time (M= 1.30, SD=1.94) and for the last time (M= 0.005, SD= 0.78) conditions $t(29) = 4.12$ $p = 0.000$ $p < .05$.

For fear in the fear video, there was a significant difference in the scores for the first time (M=0.206, SD=0.78) and for the last time (M=0.226, SD=1.07) conditions $t(29) = 2.36$ $p = 0.025$ $p < .05$.

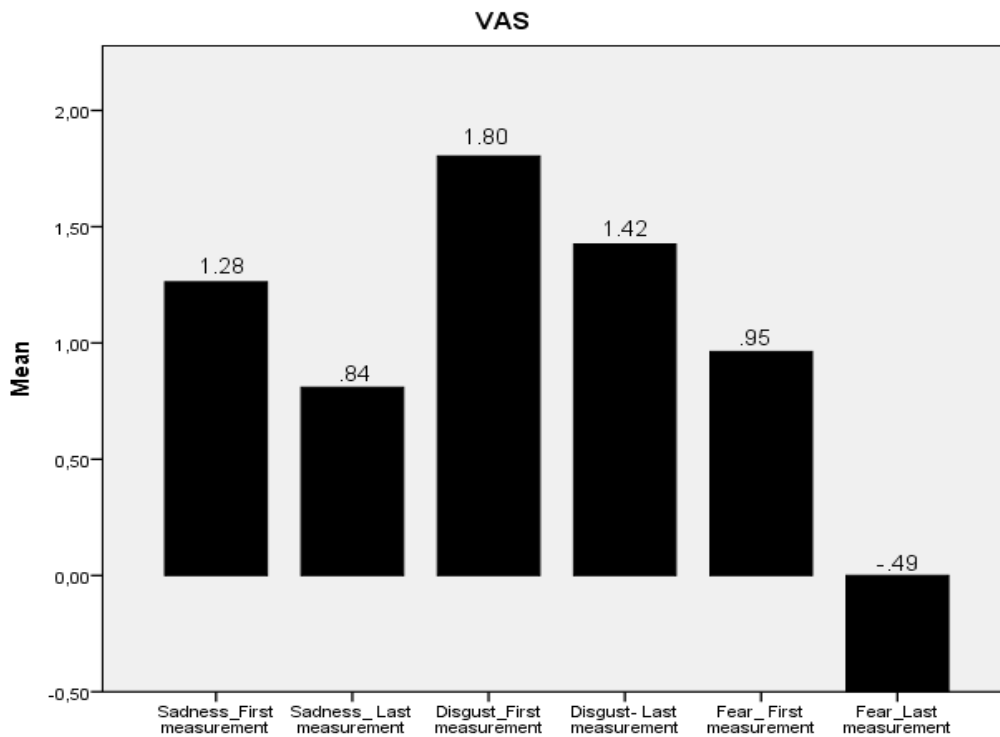


VAS- First Measurement vs Last Measurement

Two-tailed paired samples t- tests were conducted to compare the measurement of VAS in the first-time point and in the last time point for the sadness emotion in the sad evoking video for disgust in the disgust video and for fear in the fear video. For sadness in sad video, there was not a significant difference in the scores for the first time (M=1.28, SD=0.74) and for the last time (M=0.84, SD=0.74) conditions $t(28) = 2.29$ $p=0.036$ $p<.05$.

For disgust in disgust video, there was not a significant difference in the scores for the first time (M=1.80, SD=0.73) and for the last time (M=1.42, SD=0.88) conditions $t(27) = 2.32$ $p=0.028$ $p<.05$.

For fear in the fear video, there was not a significant difference in the scores for the first time (M=0.95, SD=0.77) and for the last time (M=0.49, SD=0.39) conditions $t(28) = 8.74$ $p=0.000$ $p<.05$.



Helmert Contrasts

Helmert contrasts compared the mean of each level of the factor to the mean of subsequent levels. So, the output showed us if the clips evoked the emotion they were supposed to.

We conducted Helmert contrasts separately for FaceReader, EMG and VAS. Looking at the results, during the sad video, sadness was experienced the most according to measures of FaceReader, EMG and VAS. FR: $F(1, 25) = 208.51, p = .000, \eta_p^2 = .893, p < .05$; EMG: $F(1, 25) = 4.79, p = .038, \eta_p^2 = .161, p < .05$ VAS: $F(1, 25) = 39.480, p = .000, \eta_p^2 = .612, p < .05$

In the disgust video, according to VAS results, disgust was experienced the most. $F(1, 25) = 218.819, p = .000, \eta_p^2 = .897, p < .05$; But according to EMG and FaceReader, sadness was the most experienced emotion during disgust video. EMG: $F(1, 25) = 6.825, p = .015, \eta_p^2 = .214, p < .05$ FR: $F(1, 25) = 92.79, p = .000, \eta_p^2 = .788, p < .05$ In the results of fear video, we saw that sadness was the most experienced emotion during the fear video according to the FaceReader. $F(1, 25) = 269.36, p = .000, \eta_p^2 = .915, p < .05$ EMG did not report any emotion was significantly more or less experienced than the others during the fear video According to VAS

fear was the most experienced emotion during the fear video. VAS: $F(1, 25) = 39.722$, $p = .000$, $\eta_p^2 = .614$ $p < .05$

We also conducted the Helmert contrasts to the neutral video outcome. In the neutral video, sadness was experienced the most according to the measurement of FaceReader and EMG. FR: $F(1.26, 31.56) = 391.23$, $p = .000$, $\eta_p^2 = .940$ $p < .05$ EMG: $F(1, 25) = 5.54$, $p = .027$, $\eta_p^2 = .181$ $p < .05$ According to VAS, fear was experienced most in the neutral video. $F(1, 25) = 8.68$, $p = .007$, $\eta_p^2 = .258$ $p < .05$

DISCUSSION

The present study investigates the efficacy of FaceReader compared to EMG in measuring facial expressions triggered by different emotions. The other research question of this study was if there was a difference between the measurements of FaceReader, EMG and VAS for the first-time point and the last time point.

We analyzed the results in three different sections for each of our hypothesis; the results of sadness emotion in sad video, disgust emotion in disgust video and fear emotion in fear video. The main hypothesis of this thesis was “Facial expressions of an individual for the emotions; disgust, fear, sadness can be measured accurately by FaceReader and the results would be coherent compared to facial electromyography.”

This study has shown that the measurements of FaceReader and EMG on the first-time point were statistically different from each other, for sadness, disgust, and fear. The possible reasons for the unexpected findings which might have affected the efficacy and outcome of the FaceReader will be discussed in next paragraphs.

The Helmert Contrasts were conducted to clarify if the movie clips chosen for this study really evoked the emotion expected to be evoked the most during that video in participants. The outcome demonstrated that sadness was the most experienced emotion during the sad video according to FaceReader, EMG and VAS. Disgust was the most experienced emotion during the disgust video and fear was the most experienced emotion during the fear video according to VAS.

The unexpected findings of FaceReader and EMG suggested that the disgust video evoked sadness more than other emotions in participants. FaceReader suggested that sadness was experienced the most during the fear video. However, the outcome of VAS suggested that

all the videos evoked the emotions they were supposed to. In considering that VAS outcome refers to participants' actual thoughts and feelings, and FaceReader and EMG only refers to the reflection of those feelings on the participants' faces, we might assume that the most likely explanation for the inconsistent findings is that either the emotions did not reflect enough on participants' faces as facial expressions or FaceReader did not capture those emotions the most during disgust and fear videos. This leads us to limitations of this study.

The outcome suggested that FaceReader and EMG's measurements are not coherent. FaceReader's measurement of both fear and disgust were actually very close to zero, and this made it difficult to have a good comparison. In this study, we realized by examining our outcome and the existing literature, FaceReader tends to interpret neutral face as sad facial expression, which was also broadly consistent with the finding of Chia-Yin Yu and Chih-Hsiang Ko in their study "*Applying FaceReader to Recognize Consumer Emotions in Graphic Styles*" (2017). A movie clip is approximately 29 seconds long and FaceReader analyzes the whole 29 seconds to figure which emotion was felt the most during that time. If we consider FaceReader's tendency to interpret neutral face as a sad face and the possibility of there would be neutral expressions more than emotional expressions during watching a movie clip, this is the most likely explanation for the unexpected findings about sadness was felt most during disgust and fear videos and this might also be one of the reasons for the significant difference between FaceReader's and EMG's measurements.

When interpreting the findings of this study, we should keep in mind that there is not a specific muscle getting activated for each facial expression. When we examine the "*Facial EMG as a Tool for Inferring Affective States*" by Anton van Boxtel (2010), we see that the muscles getting activated and the activity produced in the face are highly similar for fear and sadness; both include raising eyebrows and lowering eyebrows. It has also been revealed in previous studies that sadness, fear, and anger are all activating the activity of corrugator muscle (brow muscle). (Hu & Wan, 2003) Hu and Wan (2003) also found in their study that negative self-generated emotions are activating corrugator muscles. This similarity between the muscle movements is likely to challenge the distinction FaceReader is supposed to do between sadness and fear. This could also be the explanation for the EMG's inconsistent finding with VAS, because EMG also analyzes the muscle movements and the sadness and fear activity in face are very similar.

If we take notice of the VAS outcome, we see that participants thought and felt that the video was disgusting. Also, we would like to mention that, it was stated by a quite a few participants in the comments about the experiment part of the questionnaire, that that specific video was very disgusting. The same video was also found to be highly evoking disgust in the previous studies. (Gross, & Levenson, 1995) It is likely that the FaceReader had a challenge in distinguishing disgust. Our findings about disgust was not recognized effectively by FaceReader are not new, several studies also found it. (Terzis, Moridis, and Economides 2010) In the EMG's measurement during the disgust video, even though it was not measured as the most felt emotion, disgust's measurement was very close to sadness. It was found by D'Arcey (2013) in his study "*Assessing the validity of FaceReader using facial EMG*" that corrugator muscle (brow muscle) which was earlier found to be related to be sadness fear and anger was also significantly related with the facial expression caused by disgust. This means that sadness, fear, and disgust are all activating similar muscles in face when they are felt by people and this could be a possible explanation for the inconsistency of the measurement of these emotions.

To our knowledge, in contrast to most of the studies in the literature testing the FaceReader (Bența et al., 2009; Lewinski et al., 2014), movie clips were used in this study to trigger emotions instead of pictures This study aimed to measure the average of spontaneous facial reactions instead of the peak values caused by still images. It has also been discussed in previous studies that analyzing facial expressions through videos could be superior compared to analyzing through still images. (Dacu & Rothkrantz, 2007). FaceReader had 29 seconds to analyze for each video and Face Reader records 15 frames per second (Noldus, 2017). A lot of frames of faces were analyzed especially compared to the studies analyzed facial expressions of participants by showing them pictures. During watching a video, neutral expressions were done by people in the beginning and when they are trying to understand what is going on and also when they watch it again. Even though the participants show the emotional expression expected in the peak points of the movie clip, those expressions might be less than the neutral expressions. In considering the FaceReader's tendency to measure neutral as sadness, 29 seconds long videos might be too long for FaceReader and this might explain the inconsistent findings about fear and disgust videos. But on the other hand, an amount of time is needed for people to understand a story and get their emotions triggered by it. This could also be a limitation of FaceReader and might be another possible reason of the unexpected finding of the comparison between FaceReader and EMG. Also, if we want to use FaceReader in clinical

practice, we should remember that FaceReader is going to be used for longer videos. So, more research about this issue is needed.

The second hypothesis of this thesis was “There is a significant difference between the facial reactions of people between the first time and the fifth time that they watch a movie clip which is triggering their emotions.”

The results have shown that, according to the measurements of FaceReader, there was not a statistically significant difference in the facial expressions of the participants between the first time and the fifth time they have watched the movie clips for sadness, disgust, and fear. Facial EMG suggested also that there was not a significant difference in the facial expressions of the participants for the sadness emotion between the first time and the fifth time.

Previous studies in the literature have found that repeatedly being exposed to a disgusting stimulus does not always cause a decline in disgust. (Olatunji, Wolitzky-Taylor, Willems, Lohr, & Armstrong, 2009). By taking notice of the FaceReader outcome, we could assume that the habituation did not occur for the disgust video if the VAS and EMG outcome did not suggest the opposite, which was the case. Although, EMG captured a much larger decrease in the disgust ratings than VAS did.

The VAS outcome showed that there was a significant difference between the first time and last time measurements of sadness, disgust, and fear. Facial EMG and VAS both captured that there was a statistically significant difference in the facial expressions of the participants between the first time and the fifth time for the disgust emotion when they watched the disgust video and for the fear emotion when they have watched the fear video. But the unexpected result was that EMG suggested that there was an increase in fear over time which was the complete opposite result of VAS. VAS in fact, captured the largest decrease of emotion for the fear video. Since the VAS outcome reflected the thoughts and feelings of people directly, we could assume that the habituation effect occurred and there had been a decrease in participants' emotions over time. The VAS finding about fear is consistent with the Pavlovian fear conditioning principles, which suggests repeated exposure to the exact same fear stimulus causes the fear to decrease (Fanselow & Sterlace, 2014).

Even though FaceReader suggested that there was no statistical difference of the sadness ratings between the first time and last time, the ratings of sadness were actually increased according to FaceReader, which is in contrast with the VAS outcome. The tendency of FaceReader to interpret neutral expression as sadness expression could be one of the reasons of this inconsistency, because watching the same video for the fifth time could cause neutral

expressions. In other respects, it has been discussed in the literature that neutral expressions in hold emotional significance too. (Carjaval et al., 2013) Neutral expression might be caused by boredom or confusion as well. FaceReader does not recognize boredom as an expression but it is likely that people experienced a lack of interest during the last videos since they already watched them for multiple times. Facial EMG also suggested that there was no difference for sadness and the reason might be that the boredom caused by watching the same video for the fifth time reflected on participants' faces could have been interpreted as sadness. We saw that most of the participants started to lean in their chair and looking around or looking down as the time goes by, which could be counted as body signs of boredom (Kroes, 2005). When participants leaned in their chairs, a part of their faces got out of the camera angle in some cases. Even though we asked the participants kindly to sit straight before the experiment, it was not possible to remind it to them during the experiment without causing distraction. This problem could have affected the outcome of FaceReader.

During the whole experiment, participants had electrodes of Facial EMG on their faces. Those electrodes covered a part of the face and this might have influenced negatively the camera angle of FaceReader. Also, it is possible that participants felt stressed caused by having electrodes placed on their faces and felt worried about the electrodes might drop from their faces. Having an observer in the room might have also caused stress for participants. It was shown by previous studies that people tend to modify their behavior when they are being observed. (Burnham, 2007) These should also be taken in consider as factors which might have caused the participants to show less facial expressions than they felt.

Measuring emotions is not a simple process and there are multiple variables which constitutes emotions, therefore it is difficult to have a standard emotion measurement tool. (Mauss& Robinson, 2009) If we approach the research question "How can we measure emotions?" from a different angle, we could also suggest that emotions might need some more resources than facial expressions to be measured. It has been studied earlier in the literature that facial expressions could be insufficient in measuring emotions and the tone of voice and the autonomic nervous system reactions are also showing the affects when people show little emotion in their facial reactions (Barrett, 2006). So, doing some measurements of the tone of voice or autonomic nervous system reactions of people in addition to measuring facial reactions could be a good improvement for the future emotion measurement studies.

This study was done to research the utility of FaceReader. Even though we could not provide evidence by our findings about the efficacy of FaceReader in this study, this does not enable us to think that FaceReader could measure facial reactions effectively. We see the potential in this measurement instrument but future studies and implications are needed before being able to use FaceReader in clinical practice.

Several problems remain to be solved in future researches before suggesting that FaceReader could be used in clinical practice, for instance; the confusion between sadness and fear measurements of FaceReader, the confusion of FaceReader between neutral expression and sadness expression, the problem about distinguishing disgust, FaceReader's lack of accuracy in longer videos.

REFERENCES

- Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on psychological science*, 1(1), 28-58.
- Beñta, K.I., van Kuilenburg, H. Eligio, U.X., den Uyl, M., Cremene, M., Hoszu, A. & Creț, O. (2009). *Evaluation of a System for RealTime Valence Assessment of Spontaneous Facial Expressions*
- Burnham, T. C., & Hare, B. (2007). Engineering human cooperation. *Human nature*, 18(2), 88-108.
- Carvajal, F., Rubio, S., Serrano, J. M., Ríos-Lago, M., Alvarez-Linera, J., Pacheco, L., & Martín, P. (2013). Is a neutral expression also a neutral stimulus? A study with functional magnetic resonance. *Experimental brain research*, 228(4), 467-479.
- Charlesworth, W. R., & Kreutzer, M. A. (1973). Facial expressions of infants and children. *Darwin and facial expression: A century of research in review*
- Darwin, C. (1872). *The expression of the emotions in man and animals*.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*
- Datcu, D., & Rothkrantz, L. (2007). Facial expression recognition in still pictures and videos using active appearance models: a comparison approach. In *Proceedings of the 2007 international conference on Computer systems and technologies*
- D'Arcey, J. T. (2013). Assessing the validity of FaceReader using facial EMG.
- Ehring, T., & Quack, D. (2010). Emotion regulation difficulties in trauma survivors: The role of trauma type and PTSD symptom severity. *Behavior therapy*
- Ekman, P. (1993). Facial expression and emotion. *American psychologist*
- Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: Investigator's Guide*. Consulting Psychologists Press.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the Human Face: Guide-lines for Research and an Integration of Findings: Guidelines for Research and an Integration of Findings*. Pergamon.
- Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of communication*

- Fanselow, M. S., & Sterlace, S. R. (2014). Pavlovian Fear Conditioning. *The Wiley Blackwell Handbook of Operant and Classical Conditioning*, 117.
- Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & emotion*, 9(1), 87-108.
- Hu, S., & Wan, H. (2003). Imagined events with specific emotional valence produce specific patterns of facial EMG activity. *Perceptual and Motor Skills*
- Izard, C. E. (1971). The face of emotion.
- Kroes, S. (2005). Detecting Boredom in Meetings. *University of Twente*, 1-5.
- Lewinski, P., den Uyl, T. M., & Butler, C. (2014). Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics*
- Matsumoto, D., & Hwang, H. S. (2011). Judgments of facial expressions of emotion in profile. *Emotion*
- Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M. A. U. R. E. E. N., & Frank, M. (2008). Facial expressions of emotion. *Handbook of emotions*
- Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and emotion*
- Montagne, Barbara, et al. "Sex differences in the perception of affective facial expressions: do men really lack emotional sensitivity?" *Cognitive Processing* 6.2 (2005)
- Moridis, C. N., Klados, M. A., Kokkinakis, I. A., Terzis, V., Economides, A. A., Karlovasitou, A., ... & Karabatakis, V. E. (2010, November). The impact of audio-visual stimulation on alpha brain oscillations: An EEG study.
- Nightingale, J., & Williams, R. M. (2000). Attitudes to emotional expression and personality in predicting post-traumatic stress disorder. *British Journal of Clinical Psychology*
- Noldus (2017). *What's new in FaceReader 7*
- Olatunji, B. O., Wolitzky-Taylor, K. B., Willems, J., Lohr, J. M., & Armstrong, T. (2009). Differential habituation of fear and disgust during repeated exposure to threat-relevant stimuli in contamination-based OCD: An analogue study. *Journal of Anxiety Disorders*

Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and Emotion*, 24(7), 1153-1172.

Stroebe, M. S., Hansson, R. O., Stroebe, W. E., & Schut, H. E. (2001). *Handbook of bereavement research: Consequences, coping, and care*. American Psychological Association.

Tomkins, S. S. (1963). Affect imagery consciousness, 2: The negative affects

Turker, K. S. (1993). Some Methodological Problems. *Physical Therapy*, 73(10/October)

Wolf, K. (2015). Measuring facial expression of emotion. *Dialogues in clinical neuroscience*

Van Boxtel, A. (2010, August). Facial EMG as a tool for inferring affective states. In *Proceedings of measuring behavior* (pp. 104-108). Wageningen: Noldus Information Technology.

Yu, C. Y., & Ko, C. H. (2017). Applying FaceReader to Recognize Consumer Emotions in Graphic Styles. *Procedia CIRP*, 60, 104-109

APPENDIX A

Informed consent*

To participate the scientifically research

“Can we face the past?”

I am informed of the research. I have read the written information (*version: 01.03.16*). I was given the opportunity to ask questions about the research. I was able to think about my participation in this research. I have the right to withdraw my permission at any moment, without given a particular reason.

I agree to participate in this research

Name:

Signature:

Date:

I, the undersigned responsible researcher, declares that the above mentioned person is orally and written informed of the above research.

Name:

Function: *Experimenter*

Signature:

Date:

- *This form is intended for research with adults who are mentally competent. In this research, the person must give permission by itself.*

APPENDIX B
QUESTIONNAIRE

Demographic data

1. Please indicate your gender:
 - Female
 - Male
 - Do not wish to answer

2. Please indicate your age o

3. What is your current occupation?
 - Working
 - Student
 - _____

4. What is your nationality?
 - _____

5. What is your native language?
 - _____

6. Did you have a dentist appointment during the past 48 hours, and if so, did your visit include any anaesthetic (narcotics)?
 - No
 - Yes
 - If so, please indicate whether you think it influenced your facial expressions in some way during the experiment
 - No
 - Yes

7. Do you have a (partial) facial paralysis or another condition that affects your facial expressions ?
 - No
 - Yes, _____

8. Are you currently taking any sort of medication that could interfere with your emotional expression?
 - No
 - Yes, _____

9. Did you ever have plastic surgery involving botox or substances with similar effects?
 - No
 - Yes, _____

Background knowledge about content of the study

1. Were you familiar with EMG before partaking in this study? [If yes, please indicate in one or two phrases what you knew about EMG]

2. Were you familiar with some of the video clips before seeing them in this study? If so, please indicate which ones:

3. Do you have any remarks or comments about any part of the study (including the video clips)?
