BODILY EXPRESSION OF EMOTIONS IN ANIMATED PEDAGOGICAL AGENTS

by

Zachary Meyer

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Computer Graphics Technology

West Lafayette, Indiana

August 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Nicoletta Adamo, Chair

School of Aeronautics and Astronautics

Dr. Bedrich Benes

School of Aeronautics and Astronautics

Dr. Christos Mousas

School of Aeronautics and Astronautics

Approved by:

Dr. Nicoletta Adamo

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The goal of this research is to identify key affective body gestures that can clearly convey four emotions, namely happy, content, bored, and frustrated, in animated characters that lack facial features. Two studies were conducted, a first to identify affective body gestures from a series of videos, and a second to validate the gestures as representative of the four emotions. Videos were created using motion capture data of four actors portraying the four targeted emotions and mapping the data to two 3D character models, one male and one female. In the first study the researcher identified body gestures that are commonly produced by individuals when they experience each of the four emotions being tested. Each body gesture was then annotated with descriptions of the movements using the FABO database. In the second study the researcher tested four sets of identified body gestures, one set for each emotion. The animated gestures were mapped to the 3D character models and 91 participants were asked to identify the emotional state conveyed by the characters through the body gestures. The participants were also asked to rate intensity, typicality, and sincerity for each emotion using a 5-point Likert scale. The study identified six gestures that were shown to have an acceptable recognition rate of at least 80% for three of the four emotions tested. Content was the only emotion which was not conveyed clearly by the identified body gestures. The gender of the character and the participants' age were found to have a significant effect on recognition rates for the emotions.

CHAPTER 1. INTRODUCTION

The chapter will both state the purpose and identify the research question to guide the research. The scope will be identified as well as the significance of this research and the data presented. Definitions of terms will be presented as they relate to computer graphics and emotions. The chapter will conclude with assumptions, limitations, and delimitations.

1.1 Statement of Purpose

Research into emotional gestures and measurement of their effects on recognition by psychologists has primarily been focused on the face, while leaving the body mostly untouched in research into emotional expression. The FACS (Facial Action Coding System) has been the primary source of facial emotional references for professional in fields such as animation and psychology (*Facial Action Coding System*, n.d.). While the face has been the focus for decades, the body and its ability to portray emotion has been relatively ignored until only recently (Karg et al., 2013).

Attempts, such as the BACS (Body Action Coding System) and BAP (Body Action and Posture Coding System), to create coding systems for body action and posture have been performed by psychologists, however research into the topic is much less compared to that of facial action (Dael et al., 2012; Huis In ’t Veld et al., 2014). In the past decade, research into body language recognition of emotion in people has been looked at the creation of autonomous systems to identify and portray emotions in virtual actors (Cheng et al., 2020; Neff et al., 2010).

The problem that this study addresses is the need for a coding system of body gestures that portray different emotions to be used in an animated agent system. Ultimately, the purpose is to identify physical gestures that portray happiness, content, frustration, and boredom in an animated pedagogical agent (APA). These emotions have been chosen due to the frequency in which they appear in educational settings. Happiness is a high energy level and positive emotion, content has a low energy level and is positive, frustration is high energy level and a negative emotion, and bored is low energy level and negative.

1.2 Research Question

RQ1: Can a coding system of body gestures be created that portray different emotions to be used in an autonomous animated pedagogical agent (APA)system?

RQ2: Can viewers correctly identify the emotional state of an APA solely from the set of affective body gestures identified in the above coding system?

1.3 Scope

This research focuses on perception of emotion from body gestures in 3D animated pedagogical agents. The wide variety of uses for animated 3D virtual characters, such as in film and game, creates an expansive field of uses for information of emotional gesture expressions. The information is of use in other scenarios, such as that of film and games, but for simplification of this study, only educational situations will be analyzed. Educational scenarios are limited to virtual characters lecturing to a learner.

1.4 Significance

The cataloging of four emotions and their corresponding bodily movements, captured using actors and motion capture technology, along with a set of rules, in the form of a database, for expressing each emotion would be a valuable reference 3D animators. This database would apply to a North American audience, as testing is focused in the United States and can be used for portrayal of emotion in virtual characters, to heighten the intensity with more pronounced movements, in accordance with the exaggeration principle of animation. Further, the development of a database would allow for algorithmic generation of animated affective agents that can clearly display emotions. With a sufficient amount of data, the creation of an affective body language database, similar to that of the FACS, would act as a tool for teaching aspiring animators body language and portrayal of emotion in virtual characters with the singular use of body gestures and posture.

This study creates a basis of further research into bodily expression of emotion by identification of particular gestures that portray the four identified emotions. Variances in gestures communicate different emotions to a viewer. The separate parts of the bodies have different uses for gestures, allowing for continued research into each of these modalities.

1.5 Definitions

The following definitions are for key words that are present throughout the study:

Affect – “An immediately expressed and observed emotion” (Psychiatry & Behavior Sciences Clerkship - University of Washington School of Medicine, n.d.)

Autonomous Pedagogical Agent – Virtual characters, driven by a set of rules to create movement, that fulfil the role of a teacher

Arousal – The strength or energy level of body language present in a character, which can vary from slow/disengaged (low) to fast/engaged (high)

Body Gesture – "a movement or position of the hand, arm, body, or head that is expressive of an idea, opinion, emotion, etc." (*Definition of Gesture | Dictionary.Com*, n.d.)

Modalities – Attributes of the body used for expression of emotion

Trunk – The torso of a humanoid character

Valence – The overall positivity or negativity present in the body language in a character, which can vary from unpleasant (low) to pleasant (high)

1.6 Assumptions

The following list is of assumptions that have been identified as potential effectors of the conducted study:

1. The gestures are assumed to convey the individual emotions.
2. Actors will be truthful in their intention to perform gestures that will be prompted to them.
3. Respondents will be assumed to be honest in their identification of affective bodily gestures.
4. All respondents will accurately attempt to identify the emotional gestures presented to them.
5. Methods used for analysis of generated data of the research are appropriate to answer the research question.

1.7 Limitations

The following list is of limitations that have been identified to be an aspect of the research conducted:

1. Body gestures are not universal and have different interpretations in various cultures.
2. Virtual presentation of emotions in animated characters can possibly influence recognition of emotion in actors.
3. Different personalities of participants can affect recognized emotion.
4. Participants are required to identify affective states using body language.
5. Gestures performed are captured using motion capture.
6. Gestures performed by actors are prompted, not natural.
7. Acting skill of actors is limited to undergraduate theater students at Purdue University.
8. The effects of gender on recognition are still not fully understood.

1.8 Delimitations

The following list is of delimitations that have been identified to be an aspect of the research conducted:

1. Cultural differences in perception of emotional gestures exist among all groups.
2. Exclusion of age as a factor of recognition.

These delimitations are made to reduce the scope of the study but will be discussed in the review of the literature.

CHAPTER 2. REVIEW OF LITERATURE

Much research has been conducted on human emotion and personality. The following literature review will cover key points of research, with particular interest placed on five major themes that have emerged during the conducted review of the literature. Each section of the literature review corresponds to a major theme. In section 2.1, prior research into classification of personality will be discussed while section 2.2 discusses classification of emotion to build a framework for what is needed to be known before it can be studied. Section 2.3 addresses how emotions are expressed using the modalities present in the human body. Recognition of emotion is addressed in section 2.4, with studies regarding the factors of culture, gender, and age and their influence on recognition explained. Lastly, emotion of animated characters is discussed in section 2.5 to round off the review of relevant studies to the research.

2.1 Classification of Personality

The expression of emotion is linked to the personality of a person. Emotion and personality, while different, have a connection. The ability to classify emotions as well as personalities into groups has been a field of study that has resulted in the creation of various systems. Classification of personality has generally used the widely accepted theory of the five-factor of personality (Norman, 1963). Norman found that there are five separate recurring factors from analysis of personality. The five traits are commonly referred to Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism. This classification system has been referred to as both the OCEAN model, or the Big Five theory. Openness is identified as the imagination or curiosity of an individual, Conscientiousness is the impulse control and organization, Extraversion is the outgoingness of a person, Agreeableness is the empathy someone might feel, and Neuroticism is the lack of emotional stability someone might experience (Kelly, 2004). Openness is noteworthy in that it has been found as a factor for the ability for an individual to recognize emotions (Matsumoto et al., 2000). Variances of these five factors have led to the creation of more in-depth systems to further expand upon the information that is presented. The EMOTE system allows for the application of the Laban Movement Analysis (LMA) system's space and effort factors to create 3D character animation (Chi et al., 2000). The PAR and PARSYS systems are the work of combining the EMOTE system with the OCEAN model to further expand upon personality and allow a broader spectrum of emotional expression. Each of the OCEAN traits were identified as either being represented as either high or low, and aligning each factor with the LMA system's space, weight, time, and flow factors (Allbeck & Badler, 2002). Early research by the United States Air Force identified five personality traits out of a possible 35 that showed a strong correlation to the OCEAN model proposed by Norman in 1963 (Tupes & Christal, 1992). Further work was made by Buss and Finn to attempt a further expansion of personality classification into instrumental, affective, and cognitive traits. Each of these traits were then broken down further into both social and nonsocial factors (Buss & Finn, 1987).

2.2 Classification of Emotion

Practical classification of emotion for the purpose of animation of virtual agents has relied on Ekman's major emotions. Ekman's theory of universal emotions across cultures was discovered on tests of tribes in New Guinea, resulting in identification of fear, anger, joy, sadness, disgust, and surprise as the basic emotions found in all humans, regardless of culture (Ekman, 1980). The emotional gestures of New Guinean tribespeople were recognizable to college students in the United States, validating that American citizens were capable of recognizing distinct emotions in people of other cultures (Ekman, 1970). The emotions identified by Ekman have been the principle focus for animators to showcase emotions in virtual agents, as well as used by researchers in the fields on psychology and anthropology. Other classifications of emotions exist, such as Robert Plutchik's eight basic emotions; fear, anger, joy, sadness, acceptance, disgust, expectation, and surprise (Plutchik, 1980). Plutchik's work was to create a more detailed classification system as opposed to the one presented by Ekman. Ekman's emotion system is the more used of the two, however. More recent work in the field of emotion classification has concluded that there are only four major emotions; happiness, sadness, fear, and anger. Each of these four is attributed to core effects of either reward, punishment, or stress (Gu et al., 2019). There is a continuing trend of researchers unable to agree upon a list of basic emotions that would encompass all humans in previous decades, but more recent work trends towards an agreement of universalities in emotion (Ekman, 2016).

The measurement of emotions, independent of the classification system used, has relied on the Russel Circumplex Model to determine the arousal and valence level of emotions (Russel, 1980). The Circumplex Model measures the activation level of an individual's movements to determine the arousal level, against the positive or negative feelings presented in the valence level. This model created uses a four-quadrant graphical representation where valence is measured upon the x-axis and arousal upon the y-axis. Figure 1 illustrates the measurement on 28 emotions using valence and arousal as measurement devices. The further right on the x-axis an emotion is placed, the more positive the emotion is recognized as while the further left on the x-axis the emotion is, the more negative the emotion is recognized as. Emotions placed higher on the y-axis are recognized as having a higher energy or activation level while emotions placed lower on the y-axis are recognized as having a lower energy or activation level.

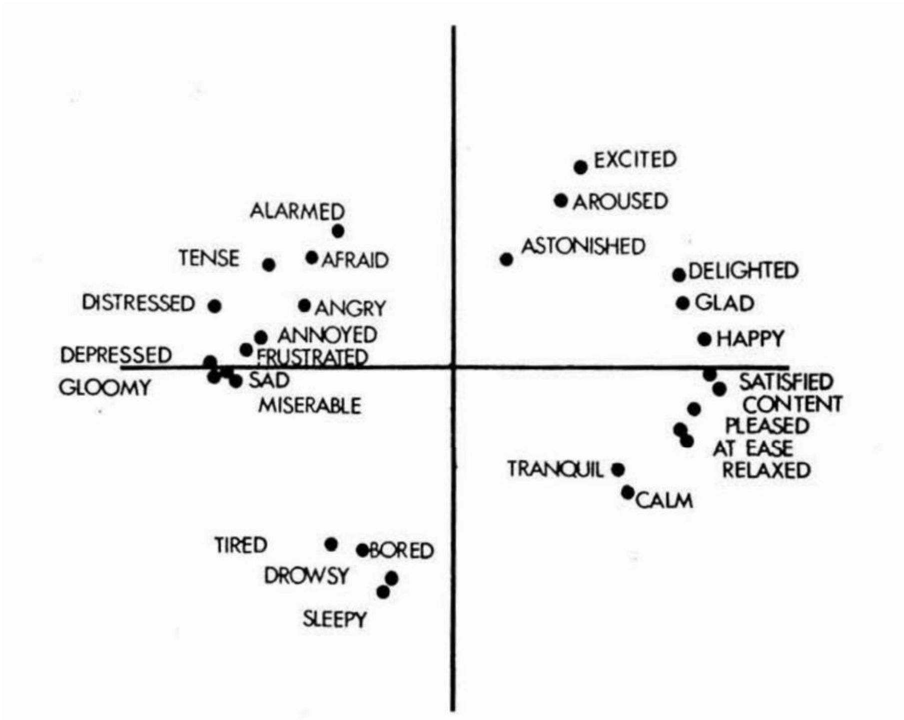

Figure 1: Russel Circumplex Model. (Russel, 1980)

2.3 Emotion Expression

While the face is capable of portraying emotion, it is limited in the axes of movement that can be used. Facial movements are a three-dimensional change, along an x-axis (left to right) and y-axis (up and down), in relation to the face, and the usage of time. The body is capable of four-dimensional expressions of movements, the x, y and z-axis (forward and backward) as well as the inclusion of time. Human emotions are expressed using several modalities: vocal and facial expression, arm and hand gestures, trunk rotation, head rotation, and leg movements. The face is cited as being the most used resource in identification of emotional state (Noroozi et al., 2018). The ability of the face to convey emotional states and the gestures associated with varying emotions have been documented in the Facial Action Coding System (FACS) (Ekman & Friesen, 1978). The work set forth by the FACS presents ways to express emotion in the face through movements of varying muscles into defined shapes. Karg et al. into facial expression of emotion has been heavily researched, but studies relevant to the body are lesser in comparison to that of facial research. A combination of modalities that include both facial and bodily gestures improves recognition rate of emotion compared to facial gestures alone by 35%. The best rate of recognition uses a combination of both gestures with the inclusion of speech (H Gunes et al., 2015). The hands are an almost equally effective modality of emotion expression as to that of the face. Hands are capable of revealing the emotional state of an individual through gestures alone. The hands themselves can be used for each of the four movement categories previously discussed. The trunk of an individual has the ability to portray a variety of emotions itself. A trunk that is contracted is recognized as sad while anger is a more expanded size and categorized with faster motion (Gross et al., 2012).

Categorization of emotion expression into databases for research purposes is an emerging trend (Karg et al., 2013). Ekman's FACS database is the primary source for facial expression of emotion, and as such, facial expression studies have generally relied on the FACS for verification. Work by psychologists towards a bimodal database of body expression and face expression has been presented to make a readily available collection of affective gestures (H. Gunes & Piccardi, 2006). The FABO database presented by Gunes and Piccardi is collection of videos that are meant to represent nine different emotional expressions and a neutral expression using a combination of the face and body as stimuli. The FABO database identifies common face and body gestures associated with Ekman's major emotions as well as three other emotions. The emotions identified in the FABO database are uncertainty, anger, surprise, fear, anxiety, happiness, disgust, boredom, and sadness. The Body Action and Posture Coding System (BAP) is a more targeted database that seeks to set forth a collection of identified actions in relation to the modalities that the body is capable producing with the intent of being used in emotion expression research (Dael et al., 2012).

The BAP database does not identify the emotions that would be associated with movements, but instead identifies the individual motions each modality of the body is capable of producing. The modalities include head orientation, head posture, trunk orientation, trunk posture, whole body posture, arm posture, gaze, head action, trunk action, and arm action. Description are present for each of the behaviors listed. Hand gestures have also been coded into databases, one of which is noteworthy. The Massey Hand Gesture Database is a collection of 2524 images of gestures readily available to the public (*Massey University*, 2012). The database is comprised of ASL gestures but does not present information pertinent to expression of emotion. The primary focus is to collect gesture into a database for use in relevant work, such as algorithms for gesture recognition.

Movements of the body can be separated into four categories according to Karg et al. to present information and communicate. The four identified categories are communicative, functional, artistic, and abstract (Karg et al., 2013). Each category is used to identify a possible purpose for each of the movements. While communicative can be used to express an emotion of an individual, a functional movement may be as simple as walking to achieve a task. Artistic movements are something you would see at a place such as a ballet performance that have the ability to express an affective state. The final category, abstract, expresses neither emotion nor is used to achieve a task.

2.4 Emotion Recognition: Effects of Culture, Gender and Age

Recognition of the emotional state of an individual has been found to have varying factors that can affect the resulting recognition of emotion. While Ekman stated that there are basic emotions that people can recognize across culture, there is further information that details that culture has an effect on recognition of emotion as well as the functional and communicative factors of gestures. Pease and Pease state that a hand gesture in one culture or country will have an completely different effect in other cultures and countries (Pease & Pease, 2004). Researchers such as Archer have identified evidence supporting that while gestures across cultures are interpreted differently, cultures also have a fascination with obscene gestures and that, while the gestures may vary, the purposes behind them remain the same. Cultures will not have the same expressive gestures that convey the same meaning, but cultures will still have a comparable gesture to convey the same meaning (Archer, 1997).

Recent work by psychologists claim that women are better at recognizing gestures and body movements and the emotional state of the individual performing the movements (Hall & Matsumoto, 2004; Mill et al., 2009). However, psychologists are not certain if women are more emotionally expressive than men (Deng et al., 2016). Researchers found that men use the right hemisphere of the brain for recognition of negative emotions, while women use the left hemisphere of the brain (Einstein & Downar, 2013). The researchers noted that the results of the study were noted as not having been replicated at the time of publication. There is also documented difference in how gender affects the perceived valence and arousal levels of emotional gestures. Researchers have found that women are more likely to rate an emotion as a higher valence level than men, but inconclusive about the effect gender has on perceived arousal (Cheng et al., 2020). Further research into the field of gender and its effect must ultimately continue to have a deeper understanding of the factors that gender attributes to emotional recognition.

Age is also another factor that can affect the perceived emotion of an individual. Children who have a higher exposure to older people in their everyday life have a higher rate of recognition of emotion in individuals (Pollux et al., 2016). The increase of age of an individual lead to the decrease in recognition rate of emotions. Older individuals are significantly less accurate in the recognition of negative emotions such as sadness and anger (Mill et al., 2009). The primary modalities tested for emotion recognition in relation to age are the vocal and facial expressions. Younger individuals were more capable of recognition of emotion using vocal modalities as a stimulus, but no significant difference was found across all ages using a facial modality stimulus (Isaacowitz et al., 2007). The Isaacowitz study also found that individuals were capable of identifying fear, sadness, and a neutral expression, regardless of age group, using vocal stimuli, but the recognition of emotion of anger, disgust, happiness, and surprise were recognized more using facial stimuli. This leads to believe that the modalities present in an individual have varying effects on the recognition of certain emotions. Facial movements and vocal expression are better at expressing different emotions in comparison to one another.

2.5 Emotion in Animated Characters

The use of animated characters for studies into recognition and expression of emotion is useful as it allows for a new testing instrument for use in studies as opposed to the use of people. In a study to identify differences between using people as the stimuli or animated characters, work by researchers found that stylized characters are more likely to have the emotion recognized as compared to realistic characters (Cissell, 2013). The study by Cissel also claims that there is little difference between body style of a character being either stylized or realistic. This allows for testing of animated body language using information and databases used in emotion research of realistic bodies. A study conducted by Noel et al found several important findings related to animated facial expressions. A static or dynamic virtual face is capable of expressing happiness, sadness, and surprise in a recognizable way to a viewer. Disgust was misidentified the most in comparison to the other emotions present in Ekman’s classification system (Noel et al., 2006). This is supported by previous findings that disgust is less likely to be identified accurately by viewers in an animated character (Fabri et al., 2004). These studies identify that an animated character is capable of expressing the same emotions that a real person would be capable of presenting to viewers. Usage of virtual characters is then a viable method for emotion recognition and expression research.

2.6 Conclusion

The literature review has presented ideas relevant to the understanding of emotion research and findings from those studies. The OCEAN model was presented as the widely accepted theory of personality characteristics. Classifications of emotions has been presented, with primary focus being placed on the Ekman classification system for identification of major emotions. Finally, Russell's Circumplex model of affect was presented as a way to measure emotions in relation to valence and arousal.

The ability of modalities to express emotion was presented, with findings detailing the ability of the primary modalities of the human body to express an emotion. The modalities as they relate to this study are arm gestures, trunk gestures, and head gestures.

Recognition of emotion research was identified, with focus being placed on factors that have the ability to influence recognition. The first, culture, shows that gestures play a role in interpretation and meaning. While several emotions are universal as has been identified by Ekman, gestures are not and can influence a viewer's ability to identify affective state. Gender was shown as a strong factor in recognition of valence and arousal levels of emotion recognition. Age was the final factor of recognition reviewed and was shown to have a decrease in recognition over time. Younger viewers are skilled in recognition, but exposure to older individuals plays a part in such.

The final section of the review briefly explores the use of animated characters as useful tools of emotion expression. Findings indicate that using animated characters show little to no difference in relation to that of realistic characters. This serves as confirmation that using an animated character is an acceptable tool for research into emotion expression and recognition of characters.

CHAPTER 3. METHODOLOGY

Contrary to the early assumption that body movement only indicates emotional intensity, recent studies have shown that body movement and posture also convey emotion-specific information (Lawson et al., 2021). The goals of this research are to identify and validate body gestures that convey emotion specific information. More specifically, the objectives are to identify and validate four sets of body gestures, each one expressing one of the following four emotional states: happy, content, frustrated and bored.

The research comprises of two studies. The objective of the first is to identify those body gestures that are commonly produced when someone is experiencing each of the four individual emotions. The goal of the second study is to validate the emotion-specific information conveyed by the identified sets of body gestures. In the second study, each of the identified 4 sets of gestures were applied to an animated pedagogical agent and a group of participants were asked to recognize the emotional state of the agent solely from its body gestures. To conduct the studies, the researcher made use of motion capture and 3D character animation technology.

3.1 Study 1: Expression Identification

3.1.1 Design

The conducted study identified, and annotated gestures related to the expression of the four emotions being tested. The researcher reviewed motion capture data of actors to qualitatively determine the emotions specific gestures present in the recording. As the researcher is reviewing the recorded motion capture data as an individual, the use of participants was not necessary for the initial study.

3.1.2 Stimuli

The actors were presented with a set of four scenarios; the goal of each scenario was to prompt one of the four emotions. The actors were then asked to express the emotion using body language only, however, they were not given clear instructions on how to perform the body gestures. The performance of each actor was video captured, and motion captured using the Xsens motion capture system and Autodesk Maya. A total of four actors were recruited from the Purdue University Theater program, two male and two female. Each of the four actors portrayed the four emotions being tested three times, resulting in a total of 48 videos and 48 motion captured takes. Motion capture data was extracted from the Xsens system and mapped onto Gabriel Salas' character models David and Dana (*Dana & David*, 2020). Each character has the facial features obscured by sunglasses and a face mask. Rotation of the sunglasses and face mask were attached to head rotation of the character to maintain head gestures.

3.1.3 Procedure

Recorded motion capture data was analyzed and annotated for each of the 48 total videos created. Annotations included total time length of gestures, measured using total amount of frames, emotion recognized, forward arm gestures, downward arm gestures, upward arm gestures, outward arm gestures, inward arm gestures, forward body lean, inward body lean, trunk rotation, head rotation, forward head gesture, and inward head gesture. Each gesture, as it is defined for the purpose of this study, is comprised of head, torso and arm movements. For each emotion tested, the gestures that appeared at least three times among the four participants were used in the secondary study. In the absence of repetition of gestures among actors, the researcher selected the two best recognizable gestures for each emotion, utilizing the FABO database as a reference to ensure they are appropriate gestures for testing.

The outcome of study 1 was an identified group of body gestures that portray the four selected emotions. The identified gestures included a set of “happy body gestures”, e.g. a set of gestures that people are likely to produce when experiencing happiness; a set of “frustrated body gestures”, e.g. a set of body gestures that people are likely to produce when experiencing frustration; a set of “bored body gestures”, e.g. a set of body gestures that people are likely to produce when experiencing boredom; and a set of “content body gestures”, e.g. a set of body gestures that people are likely to produce when experiencing content.

3.2 Study 2: Expression Recognition

3.2.1 Design

The conducted study determined recognition of emotion through modalities of gestures in 3D virtual characters. Verification of the identified affective gestures was done using a quantitative approach to determine common gestures associated with each of the emotions tested. The test utilized a within-subjects design. The following variables were analyzed in this study:

Dependent Variables: Emotion Recognition, perceived typicality, perceived intensity, and perceived sincerity all measured on a 5-point Likert scale, with a value of 1 being the lowest and 5 being the highest.

Independent Variables: Body Gesture, Participant Gender (Male, Female), Agent Gender (Male, Female).

Analysis of the data produced by the initial study was used to accept or reject hypotheses determined by the research to relate to the study. Each of the hypotheses was tested using gender (male and female) as the strata for separation into two separate groups. Gesture sets for the study were made of two full-body gestures for each of the four emotions tested. The first gesture set contained the gesture produced by the majority of the actors. The second gesture set contained the gesture performed by the majority of actors but appeared less often than the first gesture. Each gesture was tested and confirmed if there was an 80% recognition rate. The hypotheses of the study were the following:

H_1: Participants are able to accurately recognize affective displays of happiness in an animated pedagogical agent who is producing the first gesture of the identified "happy body gestures" and in the absence of facial expressions and speech.

H_2: Participants are able to accurately recognize affective displays of happiness in an animated pedagogical agent who is producing the second gesture of the identified "happy body gestures" and in the absence of facial expressions and speech.

H_3: There are differences in the participants' ability to recognize the emotional state of the agent (happy) based on participants' gender.

H_4: There are differences in the participants' ability to recognize the emotional state of the agent (happy) based on the agent's gender.

H_5: Participants are able to correctly recognize affective displays of frustration in an animated pedagogical agent who is producing the first gesture of the identified "frustrated body gestures" and in the absence of facial expressions and speech.

H_6: Participants are able to correctly recognize affective displays of frustration in an animated pedagogical agent who is producing the second gesture of the identified "frustrated body gestures" and in the absence of facial expressions and speech.

H_7: There are differences in the participants' ability to recognize the emotional state of the agent (frustration) based on participants' gender.

H_8: There are differences in the participants' ability to recognize the emotional state of the agent (frustration) based on the agent's gender.

H_9: Participants are able to correctly recognize affective displays of content in an animated pedagogical agent who is producing the first gesture of the identified "content body gestures" and in the absence of facial expressions and speech.

H_{10}: Participants are able to correctly recognize affective displays of in an animated pedagogical agent who is producing the second gesture of the identified "content body gestures" and in the absence of facial expressions and speech.

H_{11}: There are differences in the participants' ability to recognize the emotional state of the agent (content) based on participants' gender.

H_{12}: There are differences in the participants' ability to recognize the emotional state of the agent (content) based on the agent's gender.

H_{13}: Participants are able to correctly recognize affective displays of boredom in an animated pedagogical agent who is producing the first gesture of the identified "bored body gestures" and in the absence of facial expressions and speech.

H_{14}: Participants are able to correctly recognize affective displays of boredom in an animated pedagogical agent who is producing the second gesture of the identified "bored body gestures" and in the absence of facial expressions and speech.

H_{15}: There are differences in the participants' ability to recognize the emotional state of the agent (bored) based on participants' gender.

H_{16}: There are differences in the participants' ability to recognize the emotional state of the agent (bored) based on the agent's gender.

3.2.2 Participants

Participants were recruited using email announcements made to the Computer Graphics Department and recruiting through personal connections. Prolific, an online tool to collect testing participants, was the primary method to obtain subjects.

3.2.3 Stimuli

The videos presented to participants were of the gesture sets that have been confirmed by the initial study. Two gestures for each agent gender and each of the four emotions resulted in a total of 16 final videos for testing by participants. The faces of the Dana and David characters were obscured for the testing as to prevent influence of facial stimuli on recognized emotion. Motion capture data recorded with female actors was attached to the Dana character, and motion capture data recorded with male actors was attached to the David character.

Figure 2: David and Dana Character Models

3.2.4 Procedure

A questionnaire was created using the Qualtrics software to test participants on recognition of emotions expressed through the identified body gestures. Using the Qualtrics software allows for ease of use as participants are freely able to complete the questionnaire without time restriction. Each participant was asked a series of questions to determine emotion being recognized by body gestures. The first question was a multiple-choice response to determine recognized emotion of the gesture. Participants were asked to rate typicality, sincerity, and intensity of each emotional gesture on a 5-point Likert scale, with definitions for the three terms preceding the question. The 5-point Likert scale used values of 1 to 5, with the larger values being rated higher. The available emotions presented in the question were the four emotions being tested: happy, content, bored, frustration, as well as an option for not being able to recognize any of the emotions being presented. Video clips were randomly ordered and allowed the ability to be replayed as many times as needed by the participants.

CHAPTER 4. RESULTS

The results section will begin with a summary of all demographic info collected. Further details such as recognition rates and the effects from various factors such as gender of actor and gender of viewer will also be presented. The ratings of intensity, typicality, and sincerity for each emotion will also be presented. The research questions being asked in this study are as follows:

RQ1: Can a coding system of body gestures be created that portray different emotions to be used in an autonomous APA system?

RQ2: Can viewers correctly identify the emotional state of an APA solely from the set of affective body gestures identified in the above coding system?

4.1 Demographics

A total of 104 surveys were collected. 13 surveys were discarded for various factors such as finishing in under 5 minutes, or not completing at least one-third of all questions. 91total surveys were used for the analysis portion of the research. 50.8% of all respondents had completed high school, with 27.3% having a bachelor's degree. 13% of respondents have acquired a master's degree. Five people indicated they have some high school experience, one had completed trade school, and a further two have a Ph.D. There was almost an even amount of male vs female participants, with 46 females and 44 males. One person identified as non-binary. Out of the 91 participants, 24 indicated that they had some experience with character animation, with the remaining 67 having no experience in animation.

4.2 Validity

The data was plotted using a binned residual plot to check for any trends that would result in the data not being useful for drawing conclusions. The binned residual plot shown as Figure 2 does not reveal any trends in the residuals for the data. This allows for testing of the collected data from participants and for meaningful conclusions to be drawn from the factors that are tested in this study.

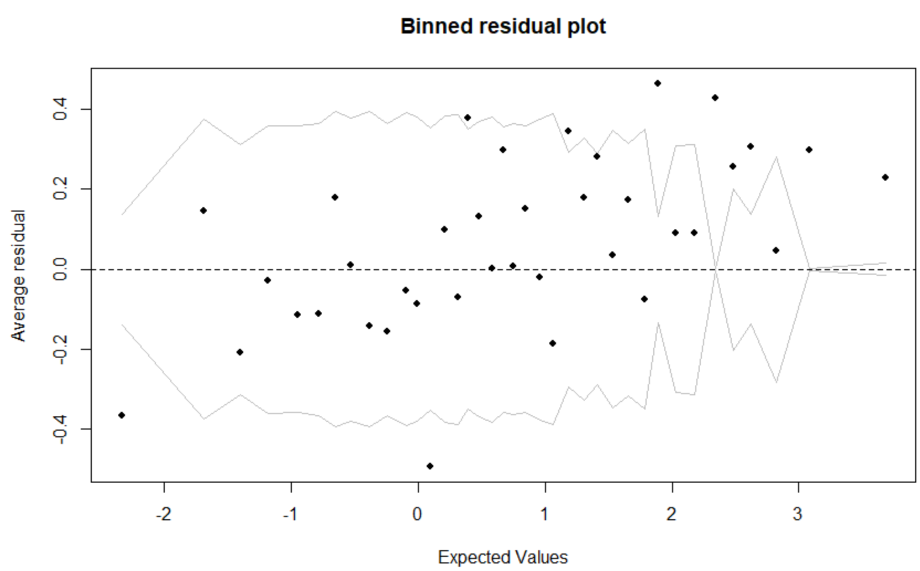

Figure 3: Binned Residual Plot

4.3 Recognition

Recognition rates of the 4 emotions varied. Participants were correct with their identification if they accurately answered the multiple-choice question with the emotion being presented to them in the video. Answers that selected “Cannot identify” were placed in the incorrect group as the emotion had not been identified correctly. A total of 1448 answers were coded for each of the 16 videos from all 91 participants. For each emotion and portrayal of emotion, body gestures were identified as being correct representations if at least 80% of participants were able to recognize the emotion. Figure 3 shows the identification count for each emotion present in the study.

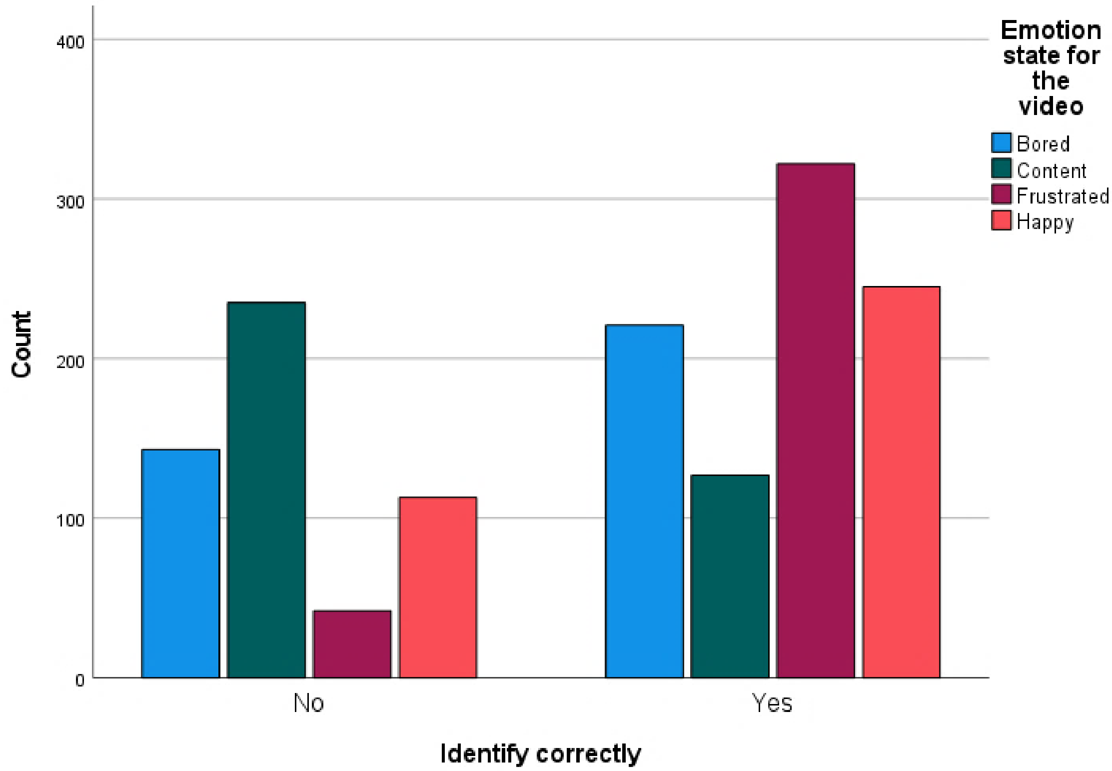

Figure 4: Recognition Rates

Of the four emotions tested, frustrated was the most commonly correctly identified emotion, with 322 responses being correct. Happy follows after with 246 correct responses. Bored was the third most commonly recognized emotion, with 221 correct responses. Content was the least correctly identified emotion, with only 126 responses having been marked as correct. Content was the only emotion that was identified incorrectly more than correctly, with a total of 238 responses being marked as incorrect. Bored and happy follow with 143 and 118 incorrect responses respectively. Frustrated had the smallest amount of incorrect responses with only 42. The recognition rate for the four emotions are as follows: 60.7% correct recognition rate for bored, 34.6% recognition rate for content, 88.4% recognition rated for frustrated, and 67.5% recognition rate for happy.

The coding system below shows the gestures for each video, with the corresponding emotional state and gender of avatar in the video. The gestures shown in Table 1 are used to identify the emotions and were validated by the participants of the study. Key videos identified as having a high recognition rate were videos 5, 6, 9, 10, 11, and 16. Table 2 gives the percentages of correct identification for all 16 videos tested. The validated videos can be accessed through this link:

https://drive.google.com/drive/folders/1SonSUKIDCHCCXTIwzKuFOta7tf-4e-8U?usp=sharing

Table 1: Video Details

Video Title	Randomized Order Number	Head Action	Trunk Action	Left Arm Action	Right Arm Action
female-bored-1	Vid 10	Downward head tilt	Forward–backward trunk leaning	Left arm action hold	Right arm action forward
female-bored-2	Vid 2	Downward head tilt	Trunk action hold	Left arm action lateral repetition	Right arm action lateral repetition
female-content-1	Vid 15	Upward head tilt	Left–right trunk leaning	Left arm action towards the body	Right arm action towards the body
female-content-2	Vid 3	Upward head tilt	Left–right trunk leaning	Left arm action towards the body	Right arm action towards the body
female-frustrated-1	Vid 9	Up-down head shake	Forward trunk lean	Left arm action forward	Right arm action forward
female-frustrated-2	Vid 5	Up-down head shake	Forward trunk lean	Left arm action forward	Right arm action hold
female-happy-1	Vid 13	Upward head tilt	Spine straightening	Left arm action forward	Right arm action forward
female-happy-2	Vid 16	Upward head tilt	Upward/forward chest movement	Left arm action upward	Right arm action upward
male-bored-1	Vid 1	Downward head tilt	Trunk action hold	Left arm action downward	Right arm action downward
male-bored-2	Vid 4	Upward head tilt	Trunk action hold	Left arm action downward	Right arm action downward
male-content-1	Vid 8	Upward head tilt	Trunk action hold	Left arm action lateral repetition	Right arm action lateral repetition
male-content-2	Vid 12	Upward head tilt	Left–right trunk leaning	Left arm action upward	Right arm action upward
male-frustrated-1	Vid 11	Head action hold	Forward trunk lean	Left arm action frontal repetition	Right arm action frontal repetition
male-frustrated-2	Vid 6	Up-down head shake	Forward trunk lean	Left arm action frontal repetition	Right arm action frontal repetition
male-happy-1	Vid 7	Upward head tilt	Upward/forward chest movement	Left arm action upward	Right arm action upward
male-happy-2	Vid 14	Upward head tilt	Upward/forward chest movement	Left arm action upward	Right arm action upward

Table 2: Video Recognition Rates

Video Number	Correct Identifications	Percentage Correct	Emotion State	Gender of Avatar
1	67/91	73.60%	Bored	Male
2	28/91	30.70%	Bored	Female
3	9/91	9.80%	Content	Female
4	51/91	56.00%	Bored	Male
5	80/91	87.90%	Frustrated	Female
6	80/91	87.90%	Frustrated	Male
7	62/91	68.10%	Happy	Male
8	23/91	25.20%	Content	Male
9	77/91	84.60%	Frustrated	Female
10	75/91	82.40%	Bored	Female
11	85/91	93.40%	Frustrated	Male
12	65/91	71.40%	Content	Male
13	37/91	40.60%	Happy	Female
14	65/91	71.40%	Happy	Male
15	29/91	31.80%	Content	Female
16	82/91	90.10%	Happy	Female

All male and female displays of frustrated were identified correctly with a recognition rate of at least 80% for all videos. A single representation of bored using a female avatar was recognized at a similar rate of recognition, as well as a single representation of happiness using a female avatar. The head, trunk, left arm action, and right arm actions for each of these physical displays can be determined to have a high enough recognition rate to validate the usage of them for emotional displays for the corresponding emotions if the recognition rate was at least 80%. Figures 4 through 9 show still frames from each of the videos that were identified as having an 80% recognition rate or higher.

Figure 5: Video 5-Frustrated gesture with up-down head movement, forward trunk lean, and forward arm actions

Figure 6: Video 6-Frustrated showing up-down head movement, forward trunk lean, and forward arm action with repetition

Figure 7: Video 9-Frustrated showing up-down head movement, forward trunk lean, and forward arm action

Figure 8: Video 10-Bored showing downward head tile, forward-backward trunk leaning, left arm action holding, and a right arm forward action

Figure 9: Video 11-Frustrated showing head action holding, forward trunk lean, and forward arm action with repetition

Figure 10: Video 16-Happy showing upward head tilt, left-right body leaning, and upward arm actions

4.3.1 Recognition of Emotion

An ANOVA test identified that there is a difference in the recognition rates between the four emotions being tested. The resultant p-value for the ANOVA test of emotions states in the videos was < .00001. This makes the recognition rates between all four emotions statistically different at a .05 significance level. It can be determined that the emotion states have different rates at which they are correctly identified by viewers. A follow-up post hoc analysis shows the differences to be in frustrated with a P-value of <.00001, content with a P-value of <.00001, and happy with a P-value of 0.048.

4.3.2 Gender of Character

An ANOVA test looking at the recognition rates of emotion based on the gender of the character showed a statistical difference in the ability to recognize emotion. With a p-value of < .00001, at a significance level of .05, the results show there is a difference in recognition rates when viewing either a male or female virtual agent produce an emotion. Male avatars had a considerably higher recognition rate compared to the female avatars. A post hoc analysis of the data shows a P-value of <.00001 for male avatars. Figure 10 shows the identification rates for male avatars vs female avatars.

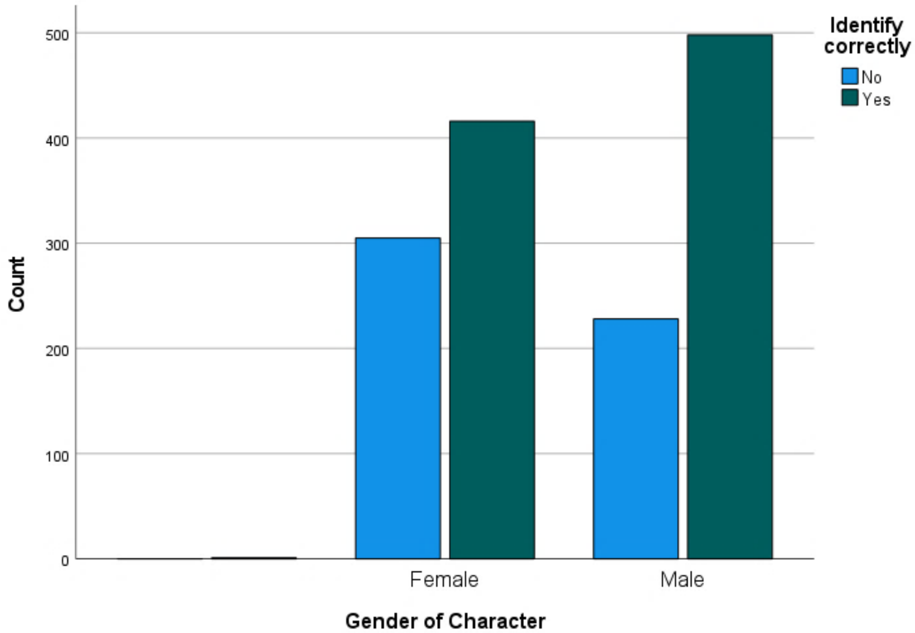

Figure 11: Identify correctly vs Gender of character

4.3.3 Gender of Viewer

An ANOVA test on the gender of the viewer identifying the emotion states of the videos showed no difference. With a p-value of 0.273, the ANOVA shows that there is no statistical difference in the recognition rates of emotions when gender of viewer is considered at a significance level of .05.

4.3.4 Animation Experience

Completion of an ANOVA test to see if people who are experienced in character animation are more likely to recognize an emotion compared to those with no experience gave a p-value was identified as 0.113. At a 0.05 significance level, it can be determined that there is no statistical difference in the ability to correctly determine an emotion based on experience as an animator.

4.3.5 Education

Testing to see if education has an effect on the ability to correctly identify an emotion was done through ANOVA. The p-value obtained from the test and was 0.826. Using a significance level of 0.05, it can be determined that the education level of an individual has no statistical difference on the ability to identify the emotion of a character.

4.3.6 Age

ANOVA testing on the age of a viewer attempting to identify the emotion of the character gives a p-value of 0.168. Using a significance level of 0.05, the data shows that there is no statistical difference in the ability to determine the emotion of a character based on the age of the person viewing the action. For each age group, the recognition rates are as follows: 16-25 years old-62%, 26-35 years old- 67%, 36-45 years old- 67%, 46-55 years old-75%, and 56+ years old-88%. Figure 11 outlines the distribution of ages vs recognition rates for the age ranges.

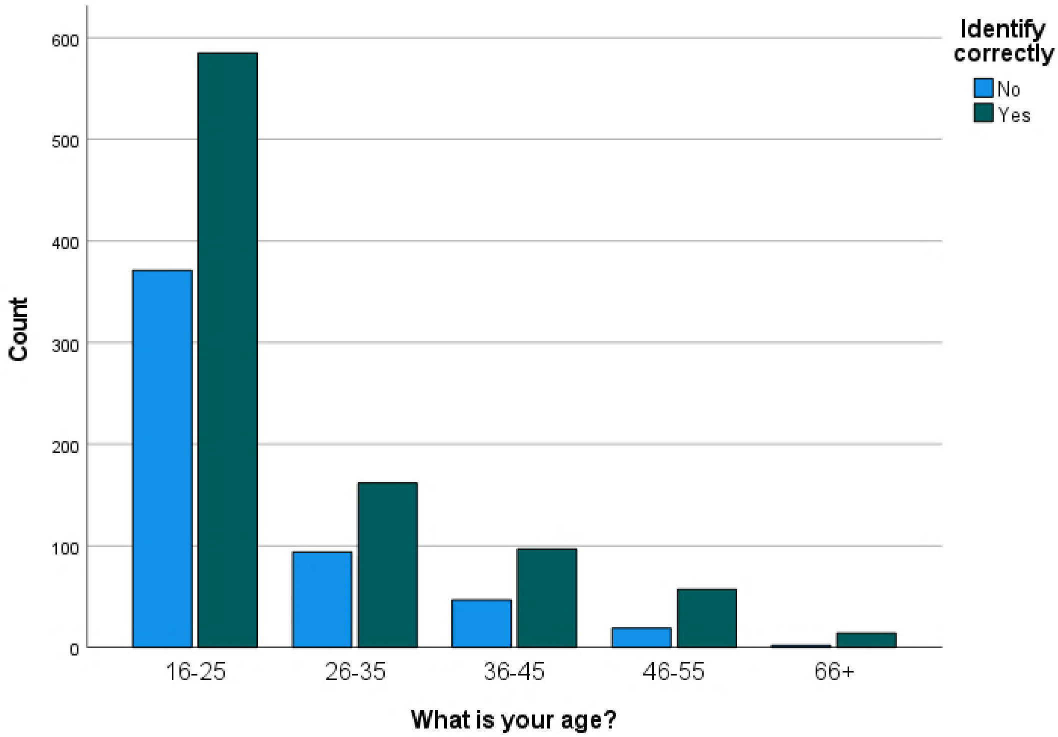

Figure 12: Recognition rates of emotion by age

4.3.7 Intensity

Intensity for each emotion was ranked on a 5-point Likert Scale. Values were between 1 and 5, with 2.5 being neutral. The average rating for the emotions is as follows: frustrated at 3.96 intensity, bored at 3.25 intensity, happy at 4.16 intensity, and content at 3.35 intensity. The intensity ratings for each video are detailed in Table 3. Figure 12 shows the recognition rates of all emotions vs intensity levels identified by participants. Figure 13 showcases the intensity rating for each of the emotions. Videos with higher intensity were found to be recognized more commonly than emotions that were rated lower on the intensity scale. Happy and frustrated were identified as having a higher average intensity rating, which aligns with the arousal levels that the emotion is classified as having.

Table 3: Intensity rating of videos

Video	Emotion	Intensity
1	Bored	3.65
2	Bored	3.53
3	Content	4.02
4	Bored	2.47
5	Frustrated	3.82
6	Frustrated	3.96
7	Happy	4.12
8	Content	2.9
9	Frustrated	3.94
10	Bored	3.38
11	Frustrated	4.15
12	Content	3.03
13	Happy	3.76
14	Happy	4.33
15	Content	3.48
16	Happy	4.44

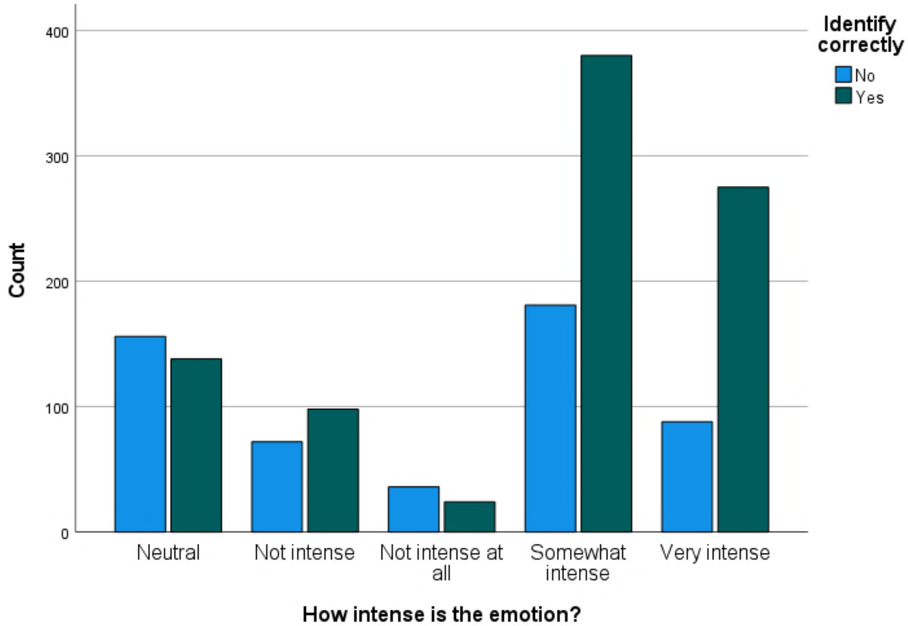

Figure 13: Identify correctly vs Intensity rating

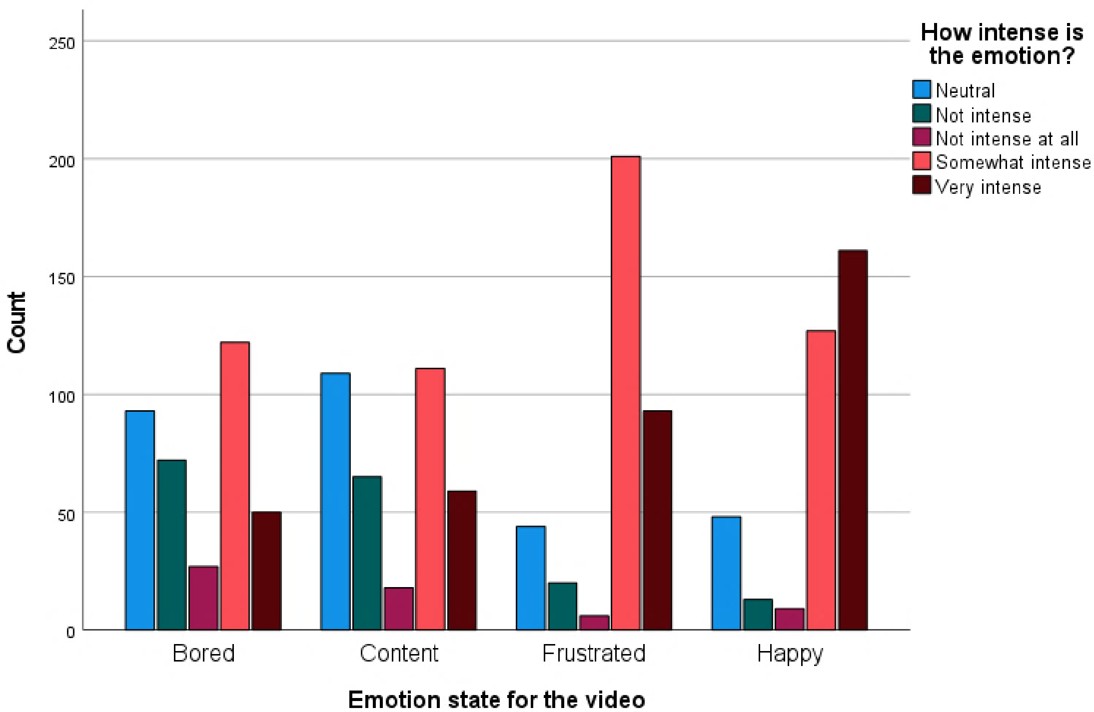

Figure 14: Intensity rating for each emotion.

Analysis of the intensity levels shows that as the perceived intensity of an emotional expression, the recognition rate for that emotion increases as well. The lower levels of intensity for each of the emotions influenced the participants in the misidentification of the emotions. Figure 14 shows the observed values for the five intensity levels along with the predicted values for intensity.

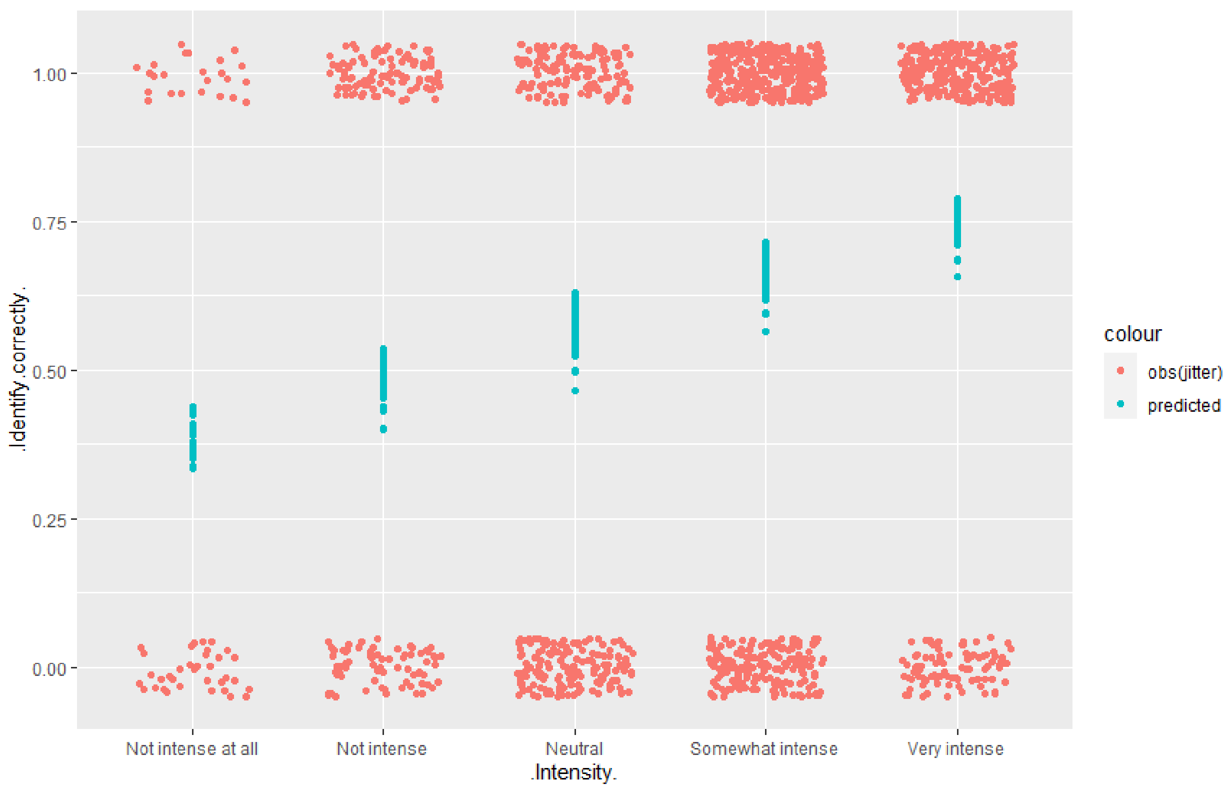

Figure 15: Observed vs Predicted values of Intensity

4.3.8 Typicality

Typicality for each emotion was ranked on a 5-point Likert Scale. Values were between 1 and 5, with 2.5 being neutral. The average rating for the emotions is as follows: frustrated at 3.94

typicality, bored at 3.72 typicality, happy at 3.34 typicality, and content at 3.29 typicality. The typicality ratings for each video are detailed in Table 4. Figure 15 shows the recognition rates of all emotions vs typicality levels identified by participants. Figure 16 showcases the typicality rating for each of the emotions. Videos with higher typicality were found to be recognized more commonly than emotions that were rated lower on the typicality scale.

TTable 4: Typicality rating of videos

Video	Emotion	Typicality
1	Bored	3.57
2	Bored	3.38
3	Content	3.03
4	Bored	3.87
5	Frustrated	3.93
6	Frustrated	4.08
7	Happy	3.53
8	Content	3.93
9	Frustrated	3.86
10	Bored	4.06
11	Frustrated	3.91
12	Content	3.32
13	Happy	3.07
14	Happy	3.38
15	Content	2.89
16	Happy	3.4

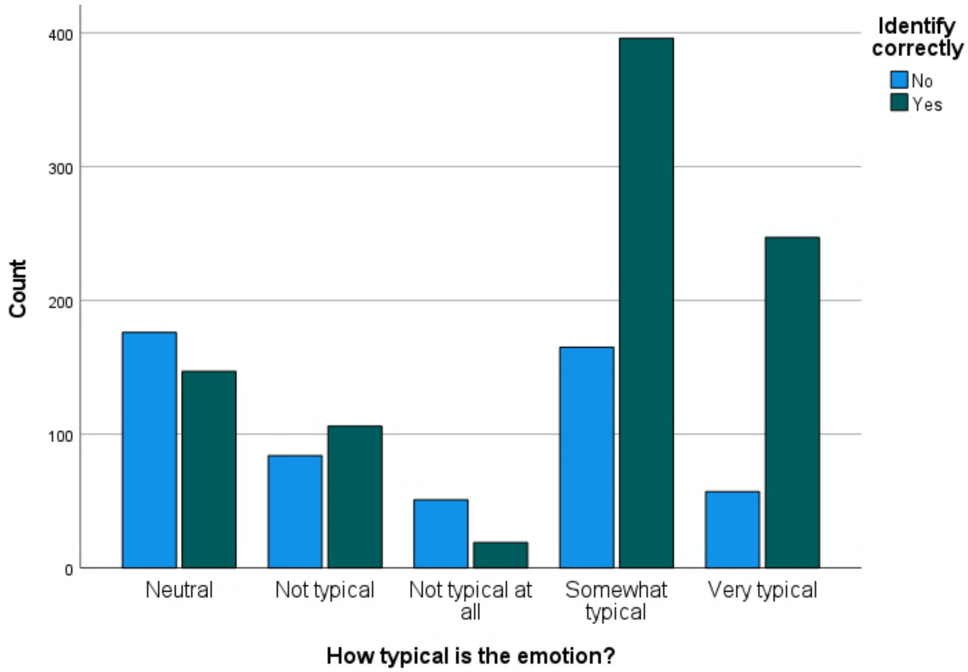

Figure 16: Identify correctly vs Typicality rating

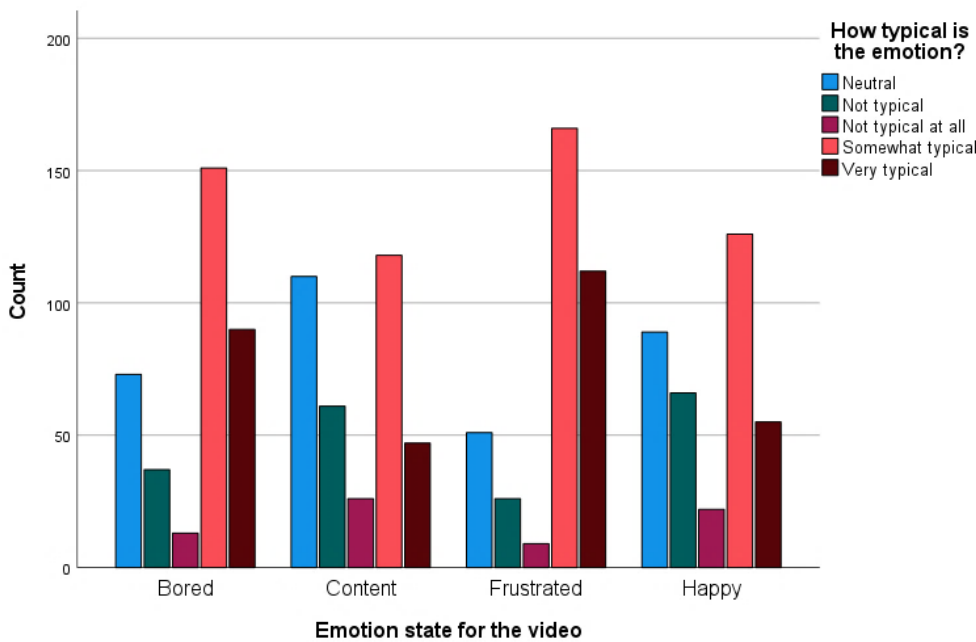

Figure 17: Typicality rating for each emotion

Analysis of the typicality levels shows that as the perceived typicality of an emotional expression, the recognition rate for that emotion increases as well. The lower levels of typicality for each of the emotions influenced the participants in the misidentification of the emotions. Figure 17 shows the observed values for the five typicality levels along with the predicted values for typicality.

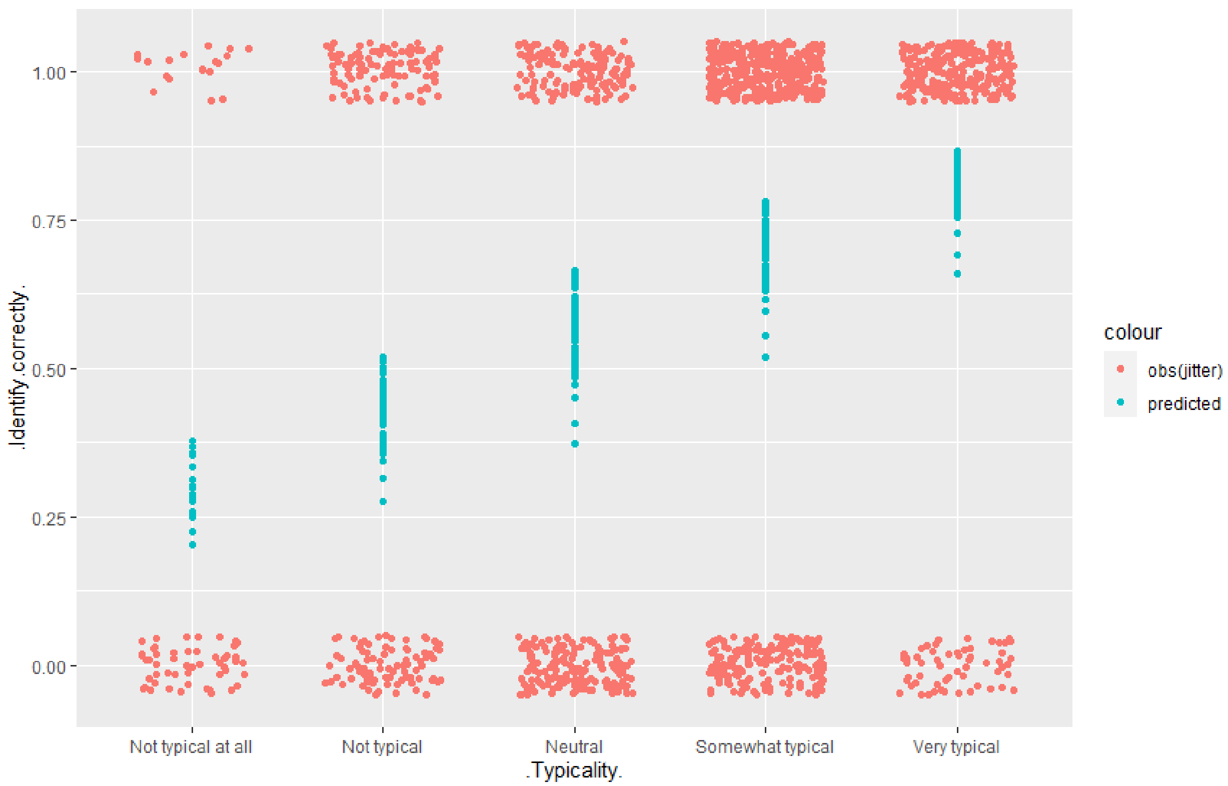

Figure 18: Observed vs Predicted values of Typicality

4.3.9 Sincerity

Sincerity for each emotion was ranked on a 5-point Likert Scale. Values were between 1 and 5, with 2.5 being neutral. The average rating for the emotions is as follows: frustrated at 4.03 sincerity, bored at 3.58 sincerity, happy at 3.63 sincerity, and content at 3.41 sincerity. The sincerity ratings for each video are detailed in Table 5. Figure 18 shows the recognition rates of all emotions vs sincerity levels identified by participants. Figure 19 showcases the sincerity rating for each of the emotions. Videos with higher sincerity were found to be recognized more commonly than emotions that were rated lower on the sincerity scale.

Table 5: Sincerity rating of videos

Video	Emotion	Sincerity
1	Bored	3.46
2	Bored	3.38
3	Content	3.2
4	Bored	3.5
5	Frustrated	4.01
6	Frustrated	4.17
7	Happy	3.78
8	Content	3.72
9	Frustrated	3.89
10	Bored	4
11	Frustrated	4.08
12	Content	3.47
13	Happy	3.2
14	Happy	3.57
15	Content	3.25
16	Happy	3.98

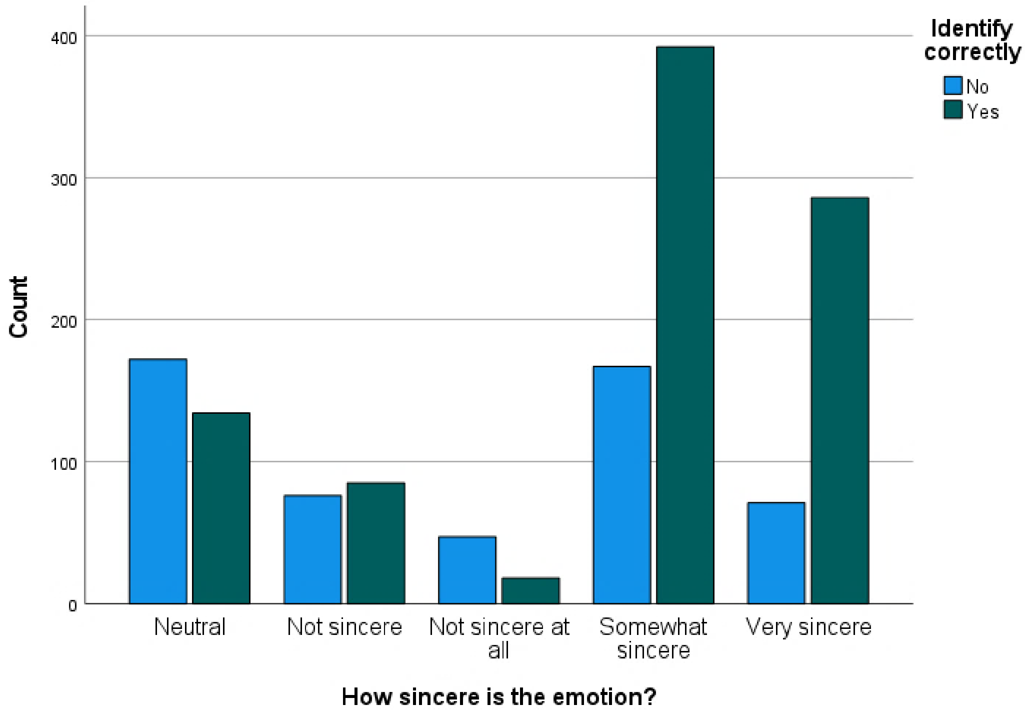

Figure 19: Identify correctly vs Sincerity rating

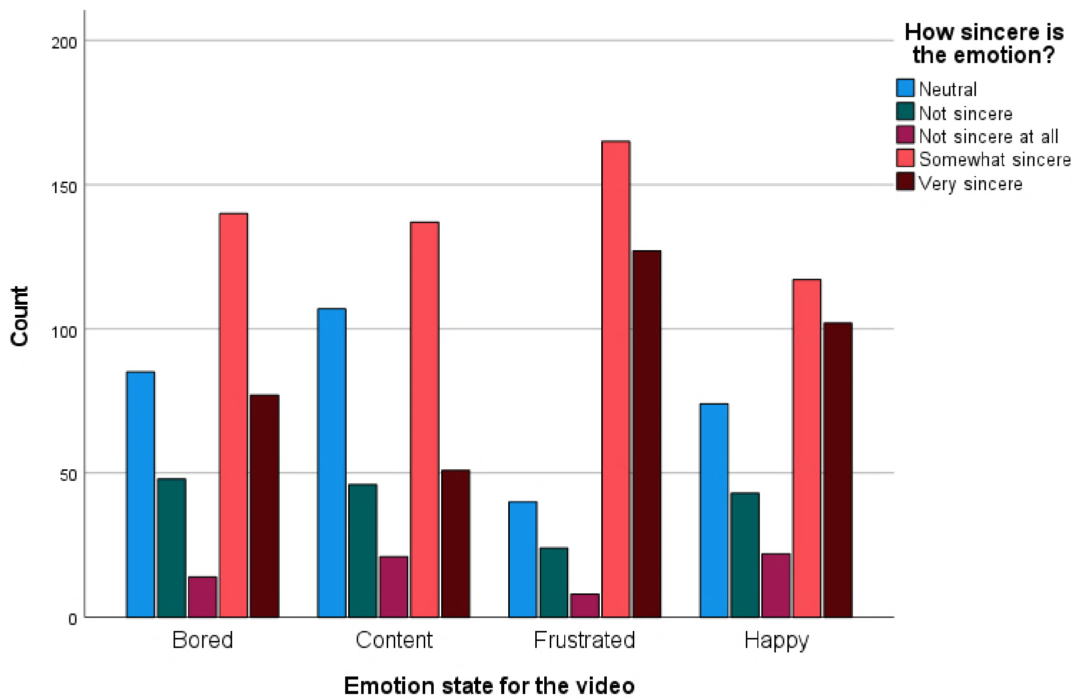

Figure 20: Sincerity rating for each emotion

Analysis of the sincerity levels shows that as the perceived sincerity of an emotional expression, the recognition rate for that emotion increases as well. The lower levels of sincerity for each of the emotions correlate to the misidentification of the emotions by the participants. Figure 17 shows the observed values for the five sincerity levels along with the predicted values for sincerity.

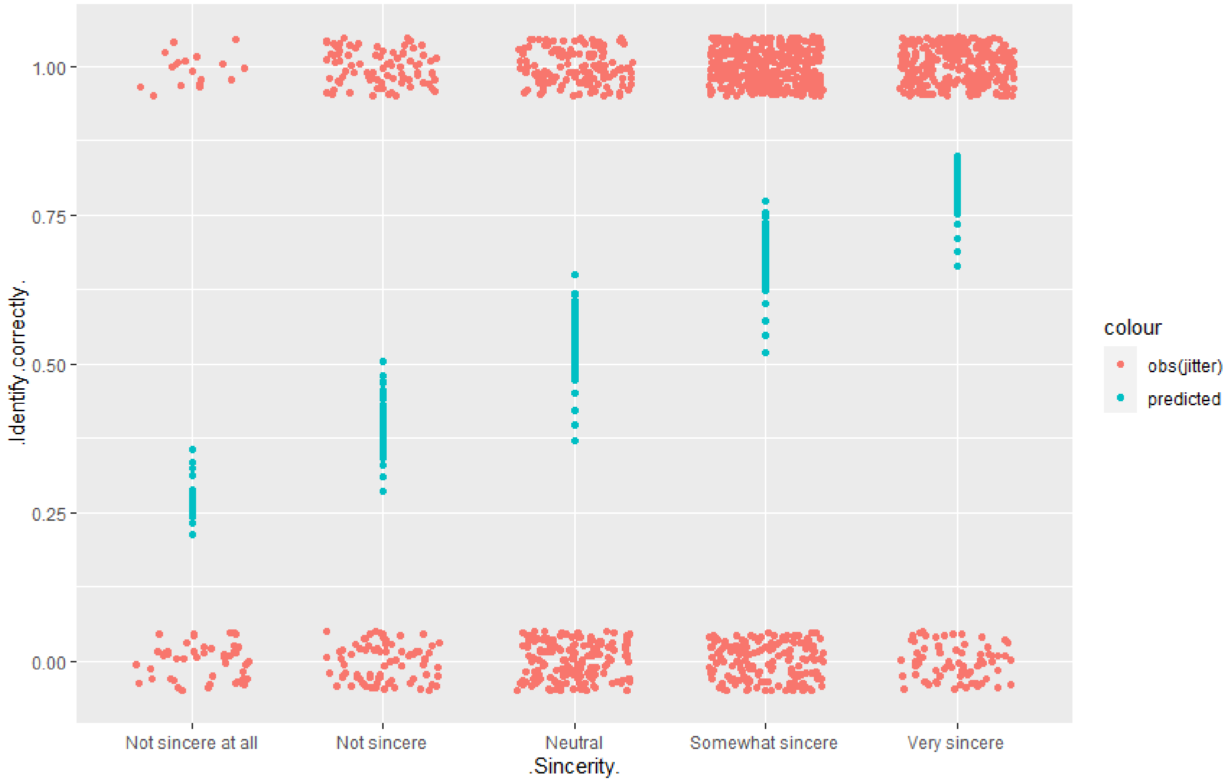

Figure 21: Observed vs Predicted values of Typicality

CHAPTER 5. DISCUSSION

The discussion section will detail the results obtained from the study. The major findings of the study are that recognition rates for the four emotions tested vary, and what gestures can be used to interpret a certain emotional state in a virtual avatar. Gender of the characters in the video also play a part in the recognition rates of the emotions shown to a viewer, as well as the age of a viewer. The following sections will detail the effect the factors have on the ability of a person to correctly recognize the emotional state of an animated character. The first section will detail recognition rates, with the effect of gender being detailed after. The contributions of age will follow after, with discussions of intensity, sincerity, and typicality finishing the discussion of the factors.

5.1 Recognition of Emotional State

A key finding identified by this study is that emotions have different rates of recognition when being expressed by a 3D avatar animated using motion capture and performing body gestures only. The easiest emotion to identify by all participants was frustrated. With a recognition rate of 88.4%, it can be said that people are highly attuned to identifying that particular emotion when conveyed by the identified body gestures. There may be varying factors as to why this may be the easiest emotion to identify of the four that were tested. It may be in part due to the fight or flight response of an individual, where a person would benefit from being able to correctly identify hostility towards them, but this would need to be tested to ensure that it is a factor in the recognition of frustration in an animated character. Bored and happy both had a higher rate of correct identification as opposed to misidentification. Bored showed a recognition rate of 60.7% while happy showed a recognition rate of 67.5%. Of the four emotions, bored and happy were the only two that did not show a statistical difference in the rates at which they were correctly identified by participants. Content was the only emotion that was misidentified more than it was correctly identified. With only a 34.6% recognition rate among participants, it shows that people are not adept at identifying content in an animated avatar with no facial features when the emotion is conveyed by the identified body gestures. It is theorized by the researcher that the lack of facial stimuli on the avatar lead to a decreased rate of recognition.

Frustrated videos shared similar motions for all, validating that the body gestures outlined in Table 1 can be used to display the emotion. An up-down shake of the head with a forward body lean, and a forward movement of the arm either once or in repetition are body movements that convey the frustrated emotional state effectively using an avatar as a means of communication. A bored emotional state can be expressed by downward head movement, with a forward body lean, and the arms having little to no movement at the sides. Avatars that show a happy emotional state can be categorized using upward head movements, with the body moving forward and upward, and the arms having upward actions. Each of these emotional displays were shown to have a recognition rate of at least 80% using the detailed actions. Whiles some of the actions had 70% recognition, for purposes of the study, 80% was the desired recognition rate to achieve for each emotional display. This is not to say that the actions that had high recognition rates were not effective in their ability to display the emotions. It is very possible that the inclusion of facial stimuli on the characters would increase the recognition rate for all videos and emotions tested, but using the body movements solely yielded results that show what actions can display an emotional state without the need for the extra stimuli.

5.2 Gender as a Factor

Data from the study shows that the avatar gender influences the emotion recognition rate. The gender of the viewer does not have any statistically significant effect on the rate of recognition for emotional state. When viewing the videos, it was shown the male avatar were correctly identified 68.8% of the time, while female avatars were correctly identified 57.2% of the time. This might possibly mean that male avatars are better at displaying emotion using body movements than female avatars. This may also be attributed to the male actors who participated in the study being more expressive in their emotional displays or the design of the characters used in the study.

5.3 Intensity, Sincerity, and Typicality

The data produced by the study shows a correlation in the recognition rates for the emotions tested in the study and the rating for intensity, sincerity, and typicality. Higher rating for intensity correlate to a higher recognition rate for each of the four emotions tested. Emotions such as happy and frustrated are much easier to identify as emotions with high intensity due to their positive arousal levels, as opposed to the negative arousal levels found in bored and content. It is possible that the extra movements associated with these emotions give extra stimuli to the viewer to help identify the emotions.

As with intensity, the data suggests a correlation in sincerity levels with the ability to recognize an emotional display. Higher levels of sincerity correlate to higher recognition rates for each of the four emotions. Being able to capture an emotion that is sincere in a controlled environment such as a motion capture studio leads to a problem such as this. The main difference, which may attribute to the high recognition rate of frustration, in particular, is that three of the four actors chose to tell a story about a time they felt frustration.

Lastly, as with the previous two factors, the typicality of the emotional gestures also influences the ability to recognize an emotion. As the perceived typicality of an emotion increases, so does the recognition rates of each emotion. The actors voiced trouble in being able to properly convey content the most as it was little movement of the body, and more focused on the face. Since the face was not included in any video, it would seem likely that displays of content were not typical.

It should be noted that there is a possibility that the three variables might be correlated with one another. Since it is possible that an emotion might have a high intensity level, the sincerity and typicality levels might be influenced by the intensity. Each of these three variables were examined individually, and not tested for correlations between them.

5.4 Implications

One implication of this research is that avatars can clearly display certain emotions using body gestures only. For instance, it may be possible for an animated agent to clearly express frustration through body cues only. Researchers and animators using 3D avatars can benefit from the knowledge that certain emotions come across better than others when displayed through a virtual avatar's body gestures.

The largest implication this research has is the beginning of a coding system of body movements that allow for the representation of emotions using virtual avatars. With frustrated being so commonly correctly identified by the participants, using the coding system of body gestures with the help of the FABO database, it can be determined that the gestures identified by the study are an appropriate method of expressing frustration in virtual avatars. Content, having been misidentified more than correctly identified, cannot follow the same conclusion as the results determined that the body gestures did not lead participants to the correct recognition of the emotion being presented. Happy and bored followed closely behind with recognition rates of 67.5% and 60.7% but were less than the recognition rate desired by the researcher. The implication of this is that the gestures for both of these emotions were inconclusive in their ability to portray the corresponding emotions, so therefore should not definitively be used for the portrayal of emotion in virtual avatars.

For usage by professional studios who create 3D character animations, they can be reassured that using the frustrated gestures identified by the study are recognized by viewers to the corresponding emotion. The gestures as having been recognized at a rate of 80% were determined to be appropriate displays of the emotions and can be used for non-verbal displays of emotion in film and games, as well as future research into the recognition of emotion in virtual avatars.

5.5 Limitations

Several limitations arose through the course of the study. First, the actors were not able to give authentic emotional displays as they had been prompted to act out the emotions through body language. Being able to capture a genuine display of the emotions while using actors wearing a motion capture suit would rely on using stimuli besides that of a prompt.

Second, a concern of the study was that people would not be truthful in their attempt to properly identify the emotion of the avatar and would finish the study without making an honest attempt. With such high recognition rates in several videos, the likelihood of that being the case seems to not be much of a concern due to the random chance of correctly identifying the emotion without making a conscious attempt.

Third, each of the 16 videos presented to the participants of the study were chosen by the researcher out of the 54 originally recorded. Since several videos were chosen due to the repetition of the physical displays between actors, there is less of a concern of bias for those videos being appropriate displays. The emotions that had little to no repetition between actors could be cases of either improper displays of the emotions, or biased choices by the researcher as to what they can identify as a proper display. As the researcher may have a different interpretation from other individuals in the recognition of emotion in the non-repetitive emotional displays, care must be taken when reaching conclusions about those individual displays.

Fourth, the study had a total of four actors. With two being male and two being female, the actors had a heavy influence on the number of videos and types of movement that would be portrayed in the study. Having more actors taking part in the study would be beneficial as the variance in movements would allow for more testing of different movements to increase the overall size of the coding system for emotional gestures. It is still important to have an equal number of actors per gender, but just an increased size of actors.

It would be beneficial to use actors that are not necessarily trained to act as this study did. Using untrained actors to portray emotion allows for more flexibility in the recruitment of participants. Since all people are capable of portraying emotion using their body movements, whether they be conscious or unconscious, it might be useful for the collection of more varied body movements.

Fifth, during the analysis of data, it was noticed that the age range for most participants who took the survey was relatively young. This heavy bias did not lead to a difference in the effect the age ranges has on the ability to correctly identify the emotion, but with such a low number of older participants, it cannot be certain that age does not have an effect. It would be much more beneficial to target equal numbers of people from all age groups to allow for a more proper analysis of age on the ability to recognize emotion in virtual avatars.

Sixth, the characters used for the study were stylized but with the facial features hidden to prevent the use of the extra stimuli as a factor of recognition. The style of the characters has a possibility to influence the recognition rate between the characters as one might seem more physically emotional compared to the other. This could be fully prevented in other studies through the use of avatars that are either simplified characters with no particular style or with no apparent gender to the character. Any of the factors that the characters style and design could cause can be rectified with the use of different characters. As part of this study was concerned with the effects of gender, the use of the stylized characters was warranted.

Seventh, of all 91 participants that were involved in the study, only one identified as non-binary on the demographic section. A sample size of one cannot have any valuable conclusions drawn from the results of statistical analysis. While the statistical analysis did not reveal any differences for the gender of the viewer, it cannot be certain if that is correct due to the singular respondent.

Eight, for each video, actors were not instructed to stand still. Several videos show an actor moving back and forth by stepping. As actors were not expressly instructed to not move back and forth, it is present in a large selection of the 16 videos. It is possible that the movement back and forth by the actors could have an influence on the recognition of the emotions being tested in the study.

Ninth, there are several times when the avatars had the hands hidden behind another body part. This was due to the camera placement remaining to the front right of each character. As it would be difficult to not hide any major body part during an expression, loss of potential visual stimuli is inevitable. The loss of stimuli over the course of the videos could have influenced the ability of the participants to correctly identify the emotion being presented.

5.6 Further Research

The research conducted can be used as a starting point for a more detailed look into the key gestures as having been identified as correct representations of emotions. While each emotional gesture was tested as a whole of all modalities combined into one full-body display of emotion, separating the gestures into their individual modalities can test if the displays are appropriate by themselves or are an individual factor of the full-body display. Further testing in this direction can lead to a more refined coding system for more gestures as well as they are tested.

The exclusion of facial features from this study was a key factor in determining if the body gestures could portray an emotion. The absence of them was also a possible source of confusion for participants as people are much more used to using the face and body combined to determine the emotional state of not only a virtual avatar, but real people as well. Inclusion of facial features for each of the 16 gestures tested in this study might reveal more insight into the need for facial stimuli when determining the emotional state of the avatars. This research might lead to more detailed information on the usage of facial stimuli in the recognition of emotion and its connection to the usage of the body as a means to display emotions.

Each emotional state was separated into male and female representations to determine if there is a difference between the avatars ability to show the emotion depending on gender. Testing should be done on whether two opposite gender avatars are capable of showing the same emotion using the same movements to determine if the recognition rates vary. This would allow for a more detailed look into if it is possible to interchange the character with the same gestures, and still retain the emotional state that is desired.

The study was limited to a total of 16 videos, and as a result, 16 gestures for four emotions. More emotions can be tested and added to the coding system, as well as the corresponding gestures for each emotion. As further research is conducted in the field of emotion recognition of body language in virtual avatars, the coding system can be further developed and used for both research as well as production of animated films and games.

This study was primarily focused on the ability to recognize an emotion from a physical gesture made by an avatar. With lower recognition rates in several videos, it appears that confusion occurred either between determining the correct emotion or not being able to recognize the emotion at all. Further research would need to be done to determine similar qualities of the physical gestures that cause the misidentification of emotions. Each video was misidentified for another emotion at least one time, but this study was not concerned with what each emotion was misidentified as. Qualitative analyses of the differences between these emotions would help with identifying more of the characteristics that each emotion uses to properly display it using physical movements.

CHAPTER 6. CONCLUSION

This purpose of this study was to identify key gestures performed by a virtual avatar to express happy, bored, content, and frustrated. The emotions were chosen due to the frequency in which they appear in educational scenarios. Qualitative analysis was performed for the first study to identify frequent actions that appeared for each of the four emotions that were chosen. A quantitative analysis of collected survey data from 91 individuals was carried out using ANOVA testing to identify factors that influence the ability to properly recognize emotions. Rates of recognition were also identified to verify the key gestures for each emotion that aid in recognition using body movements of avatars with no use of facial stimuli.

Of the four emotions tested, three emotions had key full body gestures that were identified as movements that aid in the recognition of each emotion. Content was the only emotion in the study that no gesture indicative of the emotional state could be identified properly by participants. For frustrated, body movements with an up-down shaking of the head, a forward leaning body, and a forward movement of the arms that either occurs once or in repetition is indicative of emotional state. Bored body gestures are identified as having downward head movements with a forward body lean and the arms remaining still at the side of the body. Happy body gestures were identified using upward head movements and the body moving forward and upward, with the arms moving in an upward action.

The recognition rates of the four emotions shows a strong propensity of viewers to properly identify frustration in virtual avatars without the use of facial stimuli. Bored and happy had lower recognition rates, but each had one gesture each that was properly identified at a high enough level to be considered a proper representation of the emotion. Content was the only emotion of the four that was tested that had a higher rate of misidentification.

Of the factors tested by the study, the emotional states of each video had an influence on the ability to properly recognize the emotion being shown. Different emotions had a difference in the ability to recognize the emotional state of the characters. Frustrated was the easiest to recognize among all participants, with content being the hardest to recognize. Bored and happy emotional states had a similar recognition rate among viewers.

The gender of the character was a found to have influence the ability to properly recognize an emotional state of an avatar. Male avatars had a statistically higher chance of being correctly identified in the study. This could be due to the male actors being more expressive, or the male avatar being more expressive due to the design.

The study presented identified key gestures that relate to each of the four emotions that were tested. Several factors were identified that were shown to have an effect on the ability of people to properly recognize and identify an emotional state of a virtual avatar. The creation of a coding system for emotional gestures using avatars was an important outcome of this study, as further additions can be made to the system as more gestures are identified and are added to the coding system.

REFERENCES

Allbeck, J., & Badler, N. (2002). Toward Representing Agent Behaviors Modified by Personality and Emotion. *Embodied Conversational Agents at AAMAS*, *2*, 6.

Archer, D. (1997). Unspoken Diversity: Cultural Differences in Gestures. *Qualitative Sociology*, *20*(1), 79–105. https://doi.org/10.1023/A:1024716331692

Buss, A. H., & Finn, S. E. (1987). Classification of personality traits. *Journal of Personality and Social Psychology*, *52*(2), 432–444. https://doi.org/10.1037/0022-3514.52.2.432

Cheng, J., Zhou, W., Lei, X., Adamo, N., & Benes, B. (2020). The Effects of Body Gestures and Gender on Viewer's Perception of Animated Pedagogical Agent's Emotions. In M. Kurosu (Ed.), *Human-Computer Interaction. Multimodal and Natural Interaction* (pp. 169–186). Springer International Publishing. https://doi.org/10.1007/978-3-030-49062-1_11

Chi, D., Costa, M., Zhao, L., & Badler, N. (2000). The EMOTE model for effort and shape. *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '00*, 173–182. https://doi.org/10.1145/344779.352172

Cissell, K. (2013). *A study of the effects of computer animated character body style on perception of facial expression* [M.S., Purdue University]. https://search.proquest.com/docview/1490603402/abstract/9B187DC8C8984C7DPQ/1

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). The Body Action and Posture Coding System (BAP): Development and Reliability. *Journal of Nonverbal Behavior*, *36*(2), 97–121. https://doi.org/10.1007/s10919-012-0130-0

Dana & David. (2020, March 4). https://gabrielsalas.xyz/portfolio/ava-david/

Definition of gesture | Dictionary.com. (n.d.). Www.Dictionary.Com. Retrieved November 4, 2020, from https://www.dictionary.com/browse/gesture

Deng, Y., Chang, L., Yang, M., Huo, M., & Zhou, R. (2016). Gender Differences in Emotional Response: Inconsistency between Experience and Expressivity. *PLOS ONE*, *11*(6), e0158666. https://doi.org/10.1371/journal.pone.0158666

Einstein, G., & Downar, J. (2013). Gender/sex differences in emotions. *Medicographia*, *35*(3), 271–280.

Ekman, P. (1970). Universal Facial Expressions of Emotions. *California Mental Health Digest*, *8*(4). https://www.paulekman.com/wp-content/uploads/2013/07/Universal-Facial-Expressions-of-Emotions1.pdf

Ekman, P. (1980). *The Face of Man: Expressions of Universal Emotions in a New Guinea Village*. Garland STPM Press.

Ekman, P. (2016). What Scientists Who Study Emotion Agree About. *Perspectives on Psychological Science*, *11*(1), 31–34. https://doi.org/10.1177/1745691615596992

Ekman, P., & Friesen, E. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Consulting Psychologists Press.

Fabri, M., Moore, D., & Hobbs, D. (2004). Mediating the expression of emotion in educational collaborative virtual environments: An experimental study. *Virtual Reality*, *7*(2), 66–81. https://doi.org/10.1007/s10055-003-0116-7

Facial Action Coding System. (n.d.). Paul Ekman Group. Retrieved October 28, 2020, from https://www.paulekman.com/facial-action-coding-system/

Gross, M. M., Crane, E. A., & Fredrickson, B. L. (2012). Effort-Shape and kinematic assessment of bodily expression of emotion during gait. *Human Movement Science*, *31*(1), 202–221. https://doi.org/10.1016/j.humov.2011.05.001

Gu, S., Wang, F., Patel, N. P., Bourgeois, J. A., & Huang, J. H. (2019). A Model for Basic Emotions Using Observations of Behavior in Drosophila. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00781

Gunes, H., & Piccardi, M. (2006). A Bimodal Face and Body Gesture Database for Automatic Analysis of Human Nonverbal Affective Behavior. *18th International Conference on Pattern Recognition (ICPR'06)*, *1*, 1148–1153. https://doi.org/10.1109/ICPR.2006.39

Gunes, H, Shan, C., Chen, S., & Tian, Y. (2015). Bodily Expression for Automatic Affect Recognition. In A. Konar & A. Chakraborty (Eds.), *Emotion Recognition* (pp. 343–377). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118910566.ch14

Hall, J. A., & Matsumoto, D. (2004). Gender Differences in Judgments of Multiple Emotions From Facial Expressions. *Emotion*, *4*(2), 201–206. https://doi.org/10.1037/1528-3542.4.2.201

Huis In 't Veld, E. M. J., Van Boxtel, G. J. M., & de Gelder, B. (2014). The Body Action Coding System I: Muscle activations during the perception and expression of emotion. *Social Neuroscience*, *9*(3), 249–264. https://doi.org/10.1080/17470919.2014.890668

Isaacowitz, D. M., Löckenhoff, C. E., Lane, R. D., Wright, R., Sechrest, L., Riedel, R., & Costa, P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging*, *22*(1), 147–159. https://doi.org/10.1037/0882-7974.22.1.147

Karg, M., Samadani, A.-A., Gorbet, R., Kuhnlenz, K., Hoey, J., & Kulic, D. (2013). Body Movements for Affective Expression: A Survey of Automatic Recognition and Generation. *IEEE Transactions on Affective Computing*, *4*(4), 341–359. https://doi.org/10.1109/T-AFFC.2013.29

Kelly, W. E. (2004). The “OCEAN” and the night-sky: Relations between the five-factor model of personality and noctcaelador. *College Student Journal*, *38*(3), 406.

Massey University. (2012). Gesture Dataset 2012. https://www.massey.ac.nz/~albarcza/gesture_dataset2012.html

Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., Yrizarry, N., Loewinger, S., Uchida, H., Yee, A., Amo, L., & Goh, A. (2000). A New Test to Measure Emotion Recognition Ability: Matsumoto and Ekman’s Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behavior*, *24*(3), 179–209. https://doi.org/10.1023/A:1006668120583

Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences in emotion recognition ability: A cross-sectional study. *Emotion*, *9*(5), 619–630. https://doi.org/10.1037/a0016562

Neff, M., Wang, Y., Abbott, R., & Walker, M. (2010). Evaluating the Effect of Gesture and Language on Personality Perception in Conversational Agents. In J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, & A. Safonova (Eds.), *Intelligent Virtual Agents* (Vol. 6356, pp. 222–235). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-15892-6_24

Noel, S., Dumoulin, S., Whalen, T., & Stewart, J. (2006). Recognizing Emotions on Static and Animated Avatar Faces. *2006 IEEE International Workshop on Haptic Audio Visual Environments and Their Applications (HAVE 2006)*, 99–104. https://doi.org/10.1109/HAVE.2006.283772

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, *66*(6), 574–583. https://doi.org/10.1037/h0040291

Noroozi, F., Kaminska, D., Corneanu, C., Sapinski, T., Escalera, S., & Anbarjafari, G. (2018). Survey on Emotional Body Gesture Recognition. *IEEE Transactions on Affective Computing*, 1–1. https://doi.org/10.1109/TAFFC.2018.2874986

Pease, B., & Pease, A. (2004). *The Definitive Book of Body Language*. Pease International.

Plutchik, R. (1980). A GENERAL PSYCHOEVOLUTIONARY THEORY OF EMOTION. In *Theories of Emotion* (pp. 3–33). Elsevier. https://doi.org/10.1016/B978-0-12-558701-3.50007-7

Pollux, P. M. J., Hermens, F., & Willmott, A. P. (2016). Age-congruency and contact effects in body expression recognition from point-light displays (PLD). *PeerJ*, *4*. https://doi.org/10.7717/peerj.2796

Psychiatry & Behavior Sciences Clerkship—University of Washington School of Medicine. (n.d.). Retrieved October 20, 2020, from http://depts.washington.edu/psyclerk/glossary.html

Russel, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Tupes, E. C., & Christal, R. E. (1992). Recurrent Personality Factors Based on Trait Ratings. *Journal of Personality*, *60*(2), 225–251. https://doi.org/10.1111/j.1467-6494.1992.tb00973.x

APPENDIX A: LIST OF CLIPS

Video 1: Male Bored 1	https://youtu.be/WcYJxWexxs8
Video 2: Female Bored 2	https://youtu.be/FTVVpitBG78
Video 3: Female Content 2	https://youtu.be/BghoVQWYq8Q
Video 4: Male Bored 2	https://youtu.be/rBOyevG39Xg
Video 5: Female Frustrated 2	https://youtu.be/7PWg6GkQzbM
Video 6: Male Frustrated 2	https://youtu.be/b6ElAGwdvNk
Video 7: Male Happy 1	https://youtu.be/nsEQVyQEFSs
Video 8: Male Content 1	https://youtu.be/-vgrORcqnjU
Video 9: Female Frustrated 1	https://youtu.be/p1cR2_oTtc0
Video 10: Female Bored 1	https://youtu.be/GAGw-cgq2S4
Video 11: Male Frustrated 1	https://youtu.be/10GTQIb_oQ4
Video 12: Male Content 2	https://youtu.be/6RSfS4yrBTs
Video 13: Female Happy 1	https://youtu.be/lPFneK-TYw4
Video 14: Male Happy 2	https://youtu.be/crSwoZhqHgg
Video 15: Female Content 1	https://youtu.be/ZG4b35hpl2Q
Video 16: Female Happy 2	https://youtu.be/Qf1EMdZgu-g

APPENDIX B: SURVEY QUESTIONS

What gender do you identify as?

What is your age?

What is your level of education?

Do you have experience with character animation?

Select the emotion expressed by the character. (Asked for clips 1-16)

How sincere is the emotion? (Asked for clips 1-16)

How typical is the emotion? (Asked for clips 1-16)

How intense is the emotion? (Asked for clips 1-16)

APPENDIX C: STATISTICAL ANALYSIS CODE

```
install.packages("lme4")
library("lme4")
install.packages("")
install.packages("readxl")
library("readxl")
install.packages("ggplot2")
library("ggplot2")

fulldata <- read_excel("thesis.xlsx")
fulldata$.Emotion.state.of.video. <- as.factor(fulldata$`Emotion state for the video (X1)`)
fulldata$.Subject.ID.<- as.factor(fulldata$`Subject ID`)
fulldata$.Gender.of.character <- as.factor(fulldata$`Gender of Character`)
fulldata$.Gender. <- as.factor(fulldata$`What gender do you identify as?`)
fulldata$.Experience. <- as.factor(fulldata$`Do you have experience with character animation?`)
fulldata$.Education. <- as.factor(fulldata$`What is your level of education?`)
fulldata$.Age <- as.factor(fulldata$`What is your age?`)
fulldata$.Intensity. <- as.factor(fulldata$`How intense is the emotion?`)
fulldata$.Sincerity. <- as.factor(fulldata$`How sincere is the emotion?`)
fulldata$.Typicality. <- as.factor(fulldata$`How typical is the emotion?`)
fulldata$.Identify.correctly.<- ifelse(fulldata$`Identify correctly`=="Yes",1,0)

res.aov <- aov(.Identify.correctly. ~ .Emotion.state.of.video. + .Gender.of.character
+.Gender.+.Experience. +.Education.+.Age+.Intensity.+.Sincerity.+.Typicality., data = fulldata)
summary(res.aov)
TukeyHSD(res.aov)
```

```
levels(fulldata$.Sincerity.)

since_level = c("Not sincere at all","Not sincere","Neutral" ,"Somewhat sincere" ,  "Very
sincere")

fulldata$.Sincerity. <- factor(fulldata$.Sincerity.,levels= since_level )

fin.glmer <- glmer(.Identify.correctly.~ .Sincerity. + (1|.Subject.ID.),

            data = fulldata, family = binomial)

fulldata$predicted = fitted(fin.glmer)

ggplot(fulldata,aes(.Sincerity.,.Identify.correctly.)) +

  geom_jitter(position = position_jitter(width = 0.2, height = 0.1),aes(color='obs(jitter)')) +

  geom_point(aes(y=predicted, color='predicted'))

fin.glmer <- glmer(.Identify.correctly.~ as.numeric(.Sincerity.) + (1|.Subject.ID.),

            data = fulldata, family = binomial)

fulldata$predicted = fitted(fin.glmer)

ggplot(fulldata,aes(.Sincerity.,.Identify.correctly.)) +

  geom_jitter(position = position_jitter(width = 0.3, height = 0.05),aes(color='obs(jitter)')) +

  geom_point(aes(y=predicted, color='predicted'))


levels(fulldata$.Typicality.)

typ_level = c("Not typical at all","Not typical","Neutral" ,"Somewhat typical" ,  "Very typical")

fulldata$.Typicality. <- factor(fulldata$.Typicality.,levels= typ_level )

fin.glmer2 <- glmer(.Identify.correctly.~ .Typicality. + (1|.Subject.ID.),

            data = fulldata, family = binomial)

fulldata$predicted = fitted(fin.glmer2)

ggplot(fulldata,aes(.Typicality.,.Identify.correctly.)) +

  geom_jitter(position = position_jitter(width = 0.2, height = 0.1),aes(color='obs(jitter)')) +

  geom_point(aes(y=predicted, color='predicted'))

fin.glmer2 <- glmer(.Identify.correctly.~ as.numeric(.Typicality.) + (1|.Subject.ID.),

            data = fulldata, family = binomial)

fulldata$predicted = fitted(fin.glmer2)
```

```
ggplot(fulldata,aes(.Typicality.,.Identify.correctly.)) +
  geom_jitter(position = position_jitter(width = 0.3, height = 0.05),aes(color='obs(jitter)')) +
  geom_point(aes(y=predicted, color='predicted'))


levels(fulldata$.Intensity.)
int_level = c("Not intense at all","Not intense","Neutral" ,"Somewhat intense" ,  "Very intense")
fulldata$.Intensity. <- factor(fulldata$.Intensity.,levels= int_level )
fin.glmer3 <- glmer(.Identify.correctly.~ .Intensity. + (1|.Subject.ID.),
                  data = fulldata, family = binomial)
fulldata$predicted = fitted(fin.glmer3)
ggplot(fulldata,aes(.Intensity.,.Identify.correctly.)) +
  geom_jitter(position = position_jitter(width = 0.2, height = 0.1),aes(color='obs(jitter)')) +
  geom_point(aes(y=predicted, color='predicted'))
fin.glmer3 <- glmer(.Identify.correctly.~ as.numeric(.Intensity.) + (1|.Subject.ID.),
                  data = fulldata, family = binomial)
fulldata$predicted = fitted(fin.glmer3)
ggplot(fulldata,aes(.Intensity.,.Identify.correctly.)) +
  geom_jitter(position = position_jitter(width = 0.3, height = 0.05),aes(color='obs(jitter)')) +
  geom_point(aes(y=predicted, color='predicted'))
```

APPENDIX D: SURVEY RESPONSES

Collected data from all participants can be located through the following link:

https://docs.google.com/spreadsheets/d/1etvM16ShuG6uttj0a5hiWlskh-MaHwUe67La4MaJlqo/edit?usp=sharing