Perception of Emotion Expression in Synthetic Characters

Taíssa Ribeiro
Instituto Superior Técnico
University of Lisbon
Lisbon, Portugal
taissa.ribeiro@tecnico.ulisboa.pt

***Abstract*—More and more Synthetic Characters are being used in applications worldwide. When designing Synthetic Characters that interact with human users it is important to correctly express emotions of these characters if we want to achieve more believable and effective communication. This work proposes a method based on the Repertory Grid for the creation of a model of how users perceive animations of emotions in specific Synthetic Characters, with the intent of helping development of future animations, improving the way emotions are communicated so that they are correctly recognized by the users.**

The method used two questionnaires that focused around six different animations of emotions taken from specific Synthetic Characters: Anger, Disgust, Fear, Surprise, Sadness and Happiness. Questionnaire 1 had the purpose of eliciting meaningful constructs from the participants, we selected 9 meaningful constructs through Content Analysis. As for Questionnaire 2, participants were asked to rate the animations against the selected constructs. By using Principal Component Analysis and Cluster Analysis to detect similarities we identified four clusters, where Anger and Disgust were aggregated into one cluster, then Fear and Surprise in another and finally both Sadness and Happiness were placed in their own clusters.

***Index Terms*—Synthetic Characters; Emotion Expression; Repertory Grid; Emotion Recognition.**

I. Introduction

With the emergence of more realistic synthetic characters follows the need to better communicate the emotions they are expressing. Nowadays more and more applications make use of emotion expression, either for ludic purposes as seen in video games such as The Sims 4[1], for education purposes or even health related [3], [13], [14]. Yet, emotion expression is not always correctly perceived, meaning that the emotion being expressed is not correctly identified by the receiver, which can have negative effects on the interaction.

Recognizing emotions is a difficult process for people, both with human interactions and virtual characters interactions [2], [5], [16], [21], especially without context. Even if we consider all aspects of communicating emotions, verbal and non verbal communication, confusions are still a problem.

In particular, if we consider non verbal communication, more specifically facial expressions, *how can we improve the communication of emotions in a way that they are better recognized by the users interacting with synthetic characters.*

[1]Maxis 2014, The Sims 4, Eletronic Arts

We will create a method based on the Repertory Grid. We hypothesize that with this method it is possible to create a model for how users perceive different expressions of emotions of a specific Synthetic Character. The model has the intent of detecting problems and guiding the development of future animations, improving the way emotions are communicated so that they are better recognized by the users.

II. Background

A. Emotions and Expressions

Along the years many definitions for emotion appeared and we can easily see that emotional processes and states are complex and can be analyzed from so many points of view that a complete picture is virtually impossible [15].

Even with so many definitions, for our study we will define emotions as a positive or negative experience that occur throughout life. Emotions are considered to be relatively short in duration, with changes in motor behavior, physiological changes, and cognition [12].

Ekman uses the adjective *"basic"* to list six separate emotions which differ one from another in important ways [6], Anger, Disgust, Fear, Happiness, Sadness and Surprise. We will focus on these six emotions categories because they anchor common beliefs about emotions and their expressions representing the clearest, strongest test of the common view [1].

If we want to better communicate emotions we first need to consider how they are expressed. An emotional expression is a behavior that communicates an emotional state or attitude. Each emotion expression has unique signals, the most identifiable being in the face and the voice [7]. Since we will not approach voice in our study we will focus on facial expressions, which is one of the briefest emotional signals usually lasting only mere seconds [7].

Facial expressions are *"rapid signals produced by the movements of the facial muscles, resulting in temporary changes in facial appearance, shifts in location and shape of the facial features, and temporary wrinkles"* [8]. Photographs are an efficient way to better analyse facial expressions, facial blueprints like Ekman *et al.* called it [8], and they are registered by changes in the forehead, eyebrows, eyelids, cheeks, nose, lips, and chin.

In our work, we will focus on the six basic families of emotion and its expressions: Anger, Disgust, Fear, Happiness, Sadness and Surprise.

B. Emotion Recognition

Emotion recognition represents the ability to encode an ensemble of sensory stimuli providing information about the emotional state of another individual [9].

Some common mistakes are made when trying to recognize emotions. Focusing on the facial movement patterns created by the expression of the six basic emotions, research shows that some are easily recognized, like happiness and surprise, while others are commonly confused with one another [2], [5], [16], [21]. In Bassili's research [2] we are able to see some of this confusions and where they usually happen, it being in the upper or lower areas of the face. For a brief summary refer to Table I.

TABLE I
SUMMARY OF BASSILI'S FINDINGS ON EMOTION RECOGNITION [2]. IT SHOWS COMMON EMOTIONAL CONFUSIONS BETWEEN EXPRESSED EMOTIONS (ROW) AND PERCEIVED EMOTIONS (COLUMN).

	Happpiness	Surprise	Sadness	Fear	Disgust	Anger
Happiness			x			
Surprise				x		
Sadness				x	x	
Fear	x	x	x			
Disgust				x		x
Anger				x	x	

Some other studies by Dores *et al.* [5] and Langner *et al.* [16] are able to tell us that happiness was the most frequently recognized expression, whereas anger was the least frequently recognized expression only in the former study. Both studies had a similar confusion matrix, faces with intended surprise were sometimes confused with fear, and vice versa. Intended disgust was sometimes mistaken for anger.

If we consider all of the studies above we can easily see which are the emotions that have more problems being recognized, Anger, Fear and Disgust. As seen in Bassili's research and other studies [5], [16], Anger and Disgust are a problematic case, where they are usually mistaken with each other. Although Fear is not usually mistaken for both emotions of Anger and Disgust, the same doesn't happen the other way around, even in Silva *et al.* research [21], Disgust and Anger were pointed out as being similar to Fear.

C. Repertory Grid

Repertory Grid Technique [11] is a powerful research tool in many situations, a method for eliciting personal constructs, and is normally used to explore an interviewee's views on a particular topic with the absence of researcher bias. This technique was first developed for use in psychology but throughout the years it started being used by other areas as well.

The fundamental premise on which the technique of repertory grid is based is Kelly's Personal Construct Theory (PCT) [11]. We indirectly develop *"rules"* by which we view or categorise situations, these rules are our constructs. There are two important aspects to take into account when talking about the Repertory Grid, the *elements* and the *constructs*. Kelly defined an element as *"the things or events which are abstracted by a construct"* [10] as for the latter it is *"a way in which two or more things are alike and at the same time different from one or more things"*.

The first design decision is the selection of elements. Elements should be within the range of convenience of the constructs used. Constructs applies only to a limited number of people, events or things, depending on the subject at hand. To be more specific if we are constructing a grid where the constructs will deal with the youth, having one or two old people between the elements it is not exactly in the range of convenience of the youth type of construct [10].

In order to elicit meaningful constructs, each element is written manually on a card. Different triads (a set of three elements) are presented to the interviewee until all combinations have been covered, or the interview is terminated. Five or more elements are needed to produce a sufficient number of triads so that construct elicitation can be repeated.

For eliciting constructs a question is to be asked for each interviewee when showing each triad. The general question from Kelly's work is, *"in what way are two of these alike and at the same time different from the third?"*. This will have to be adapted depending on the study at hand, but it is always important to have in mind observer bias, in other words it is important that the question does not guide the observer to a biased answer (e.g. asking "in what way are two of these better than the third?" guides an answer to be done under a value judgement).

For rating the elements of the study the constructs are reused using normally a simple five or seven point scale forming a matrix. However, other possibilities include a simple bipolar scale.

With this method we hope to achieve a better understanding of the features that are more similar between emotions and those that we consciously use to distinguish them apart. Considering that the features are our constructs and the different animations are our elements, we will have a matrix that will provide data on the features that are common in the emotions we are studying, informing us of the possible causes for similarities between them.

III. IMPLEMENTATION OF THE METHOD

Our goal is to model how people perceive expressions of emotions in an animation, so that in the future we can use it to better understand what makes those expressions similar from one emotion to another.

Our method consists in using the Repertory Grid, more specifically a Triad Analysis. We will first choose the elements for our Repertory Grid, in our case animations of facial expressions, and understand the best way to present them. Secondly, we will elicit the constructs, personal rules by which we view or categorize situations, for each individual person and finally they will be asked to rate each element against a set of selected constructs. With this method we will know the

most important constructs, in our case facial features, in each animation for each person and with the ratings understand how these are perceived by a group of people.

As Figure 1 shows our study involved two different questionnaires, one for determining the constructs by using Content Analysis which is commonly used to analyse qualitative data in Repertory Grid [4], [11], [22] and the other for the rating of the animations against each construct, these responses will be analysed using Principal Component Analysis and Cluster Analysis, again common types of analysis used in Repertory Grid [4], [11], [22]. Data gathering from both questionnaires was mainly aimed at national participants through convenience sampling. In the next sections we will explain in further detail each section of the questionnaires.

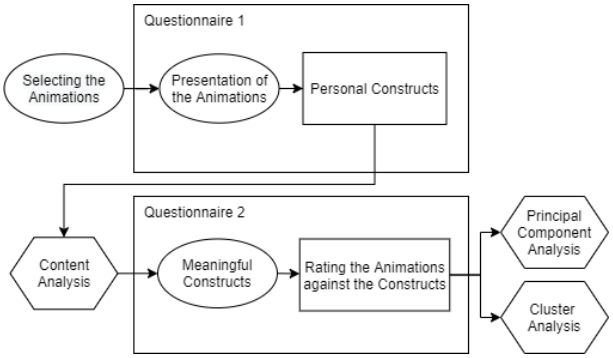

Fig. 1. Overview of the method implemented.

A. Selection of Animations of Emotions

The first step for a triad analyses is choosing the *elements* that will be presented, in our case we want to compare different expressions of emotions in Synthetic Characters. We decided on six animations, each one representing an emotion, Anger, Disgust, Fear, Happiness, Sadness and Surprise.

These animations were taken from the Virtual Tutoring Application, developed by Lima *et al.* [17] and later improved in [18], [19], [21]. In the Virtual Tutoring Application two Synthetic Characters exist, they go by the name of João and Maria, represented by a 3D head, modeled from the shoulders up, and they were provided by the modeling technology My Didimo, offered by the company Didimo, Inc [2]. An animation clip package for each of the characters was also used, developed by a team of 3D artists from the same company, and it provided the visual front end of the Synthetic Characters.

Silva *et al.* [21] provided the visual expressiveness and believability to the two virtual tutors in the application. However, the animations used were the base animations of both Synthetic Characters provided by Didimo and the team of 3D artists, without the changes from Silva's work. In Figure 2 we have a picture for each of the six animations of João in the emotion's highest intensity, Anger, Disgust, Fear, Happiness, Sadness and Surprise.

[2]Didimo, Inc, "Didimo - A digital version of you from a single photo", Software company, http://www.mydidimo.com/

Fig. 2. The six animations of João in the emotion's highest intensity. The emotions expressed are the following from top left to bottom right: Anger, Disgust, Fear, Happiness, Sadness and Surprise.

B. Presentation of the Animations

The second step is how to present the elements to the participants. The usual approach is to show to the participant three different animations, randomly presented until all combinations had been covered, or no more constructs were elicited.

To balance the amount of work required by each participant, we decided to display only a fixed number of combinations. If we take into account that we have six animations, by using the formula for combinations with no repetitions and no order restrictions we have a total of 20 combinations of 3 animations[3]. To maximize the number of different combinations seen by the participants, we decided to display six combinations per participant to allow for the elicitation of an adequate number of constructs while not tiring the participants.

To create the distribution, we considered 10 different lists each containing the six animations with different orders as seen in Table II, each letter corresponds to a certain animation of emotion, six in total A, B, C, D, E and F. If we divide the lists in two we have the 20 combinations of three animations, each list containing a combination pair.

The next step was to distribute them evenly among the participants, so that we can get the maximum amount of combinations. Since we wanted to show six combinations per person, we take three combination pairs of the ten available (e.g. pairs 0, 1, and 2 from Table II). For each participant it is provided a different set of combination pairs, more specifically if we consider n to be the number of the participant, the chosen pairs will be $n_{mod}10$, $n_{mod}10 + 1$ and $n_{mod}10 + 2$ (e.g. if $n = 13$, then the pairs chosen would be 3, 4, and 5).

Furthermore, we also decided to represent both Synthetic Characters from Virtual Tutoring Application, João and Maria, so each of the six combinations would have animations from

[3]Having 6 elements and selecting 3, the number of possible combinations equates to: ${}^{6}_{3}\mathrm{C} = 20$.

TABLE II
ALL 20 COMBINATIONS OF 3 ANIMATIONS DISTRIBUTED BETWEEN 10 PAIRS.

Combination Pairs	Combination 1	Combination 2
0	A B C	D E F
1	A B D	C E F
2	A B E	D C F
3	A B F	D E C
4	A D C	B E F
5	A E C	D B F
6	A F C	D E B
7	A D E	B C F
8	A D F	B E C
9	A E F	D B C

the different avatars. The first combination would be from João and the second from Maria, and so on and so forth. The main reason for this was to check if a different avatar would have an impact on the perception of the animations.

C. Constructs and Ratings

In this section we will explain in detail the main part of the Repertory Grid. First a small group of people will have to elicit the *constructs*, in our case facial features. Next, participants will have to rate each of the six emotions against the selected constructs.

We have to remember that the constructs used need to be meaningful to the individuals at hand, so just supplying them with no input from the participants would not work. When in doubt about what kind of constructs are applicable to a certain group of people, it is common practice to collect a sample of constructs from a comparable group or from the group itself [10]. That way you can safely assume that the most commonly used constructs for that group will be meaningful to the participants. As long as the constructs are carefully selected and meaningful to the participants, there is no reason to not use them. Therefore, we decided to split the study into two different questionnaires. Both questionnaires had some demographic questions in the beginning like the age, gender, maternal language and interaction with Synthetic Characters.

1) Questionnaire 1 - Construct Elicitation: In this questionnaire, we collect a sample of constructs from a group of people and then sort them into themes by using Content Analysis. We first give an example of how the experiment would occur, this example was given in another context with no connection to emotions so that the participant would not be biased. It mainly consisted in showing three different shapes, a circle, a square and a triangle, and we showed how we expected the answers to be.

The most common way to begin the experiment is by presenting the animations as mentioned in section III-B and then to ask a specific question, as we can see from Kelly's work [11]. Depending on the study this question may differ slightly, in our case, we first ask the participant to choose the two animations that are more alike and then to justify the choice by asking *"how are two of them alike and at the same time different from the third"*.

The questionnaire consists in six sections, each presenting a combination of three animations (see section III-B for more information in the distribution). Half of the sections present João's animations and the other half presents Maria's animations, we introduced Maria to check if a different avatar would have an impact on the perception of the animations, even though the animations are the same.

Each section had four questions:

- **Q1** *"By comparing the 3 animations presented above, identify the two that are alike"*
- **Q2** *"how are two of them alike and at the same time different from the third"*.
- **Q3** *"Provide us with one characteristic that you found was alike in the two animations"*
- **Q4** *"Provide us with the opposite characteristic from the one mentioned above, describing the different animation"*

Some notes were added in the process such as: *do not repeat characteristics between sections* and *do not use names of emotions to justify the choice of Q1*. Both **Q3** and **Q4** helped better understand the characteristics that the participants had in mind.

After all responses were received we used Content Analysis to find meaningful constructs by sorting them into themes and analysing the frequency in which they appeared. *Content analysis* is a subjective qualitative analysis by which elements and/or construct labels are placed into common categories or main issues and interpreted for meaning [4]. The selected constructs were then used to create the second questionnaire.

2) Questionnaire 2 - Animations' Ratings: After selecting an amount of constructs elicited, the ones mentioned the most, the next step in the Repertory Grid is to ask the participants to rate the animations with the selected constructs. With the Content Analysis we found 9 different constructs that were meaningful. Furthermore, the demographic characterization of the participants also included information of how many people had answered the first questionnaire, since it was not required for them to have answered it.

The questionnaire needed six sections, one for each emotion animation. The animation would be shown to the participant and, for each of the selected constructs, a seven point Likert scale was shown. The participants were asked to rate the animation using each construct presented. To guarantee that the same order of presentation of the animations would not be shown to the participants, to not create a bias on the answers, we made use of the *Latin Square*, an $n \times n$ array filled with n different symbols, each occurring exactly once in each row and exactly once in each column. In our case we will consider our six animations and we want to present them in different orders using six different versions of the questionnaire in a way that no animation occurs in the same spot in more than one version, reducing the bias that was mentioned before.

D. Tools

Throughout our study we used different tools for making all questionnaires. For Questionnaire 1 we made use of jsPsych[4]. This tool provides a flexible framework for building a wide range of laboratory-like experiments that can be run online. It also allowed us to show different videos side by side and to take the responses from previous questions into new questions. For the Questionnaire 2 we made use of Google Forms[5], it was a well known tool used for questionnaires. It provided a functionality that allows the participant to go back to previous sections in the questionnaire, something that was not trivial to make in jsPsych.

Since we were using the Virtual Tutoring Application's agents and it was developed using the Unity 3D Engine[6], we also made use of the engine to visualize the animations that Virtual Tutoring Application provided and to record them using the computer so that we could show the videos in the questionnaires.

E. Summary

This section explains how our study was conducted. We start by briefly explaining our method based on the Repertory Grid, going into the details of each important step, such as the chosen animations, the way we presented them to the participants, the constructs that were elicited and finally the rating of each animation. Furthermore, we explain the use of two different questionnaires, the first one the elicitation of the constructs, and the second one the ratings for each different animation using the constructs that were selected with the Content Analysis. We close the section by showcasing what tools were used to develop the questionnaires.

IV. Evaluation

Throughout this section, we will present and explain the results obtained for the two questionnaires that were implemented. We analysed a total of 21 responses from Questionnaire 1 - Elicitation of the constructs - and 40 responses from Questionnaire 2 - Rating of the Animations.

A. Elicitation of the Constructs

This section will describe the results of Questionnaire 1. This questionnaire allows for the collection of constructs that are understood by the participants and surrounding community, thus providing meaningful constructs to the Questionnaire 2 for rating of the animations.

We tried to find a small sample with a wide demographic spectrum. Our group of participants is aged between 21 and 56 years with a mean of 26 years old and a standard deviation of 9.540. Regarding gender 52% are male and 48% are female. As for experience with Synthetic Characters, 85.7% interacted with Non-Player Characters in videogames and 90.5% interacted with Virtual Agents. As for the interaction with Social Robots only 28.6% had interacted with one previously and finally of the 21 participants only 1 had worked with or developed Synthetic Characters/Virtual Agents.

Each participant supplied us with 6 different constructs making it a total of **126**. We had to discard a total of 25 constructs (leaving a total of 101) due to the following problems:

- We identified people that used the same construct more than once, 4 constructs were identified as having this issue;
- The use of emotions to describe the different constructs, a total of 14 to be exact, we did try to minimize this number by specifically saying not use emotions;
- There were 7 answers that had to be discarded as they weren't specific enough to get any conclusions (e.g. "eyebrows" – without an explanation of the movement or context).

The next step consists in doing Content Analysis, we took the remaining 101 constructs and found common themes in the constructs. After conducting the analyses we reduced it to 32 constructs in total. We came to the conclusion that a frequency bellow 3 occurrences was not meaningful enough to consider, given that, we can say that a total of 9 constructs were found as shown in Table III. With these constructs we reduce the possibility of not supplying meaningful constructs to complete the Repertory Grid.

TABLE III
Frequency for each meaningful construct found mentioned by the participants, sorted by most frequent to less.

Construct		
Pole 1	**Pole 2**	**Frequency**
mouth opens wide	mouth barely opens/closes	15
eyebrows go down	eyebrows go up	12
eyes open	eyes close	9
forehead goes down	forehead goes up	7
teeth visible	teeth not visible	7
face contracts	face expands	6
cheeks go down	cheeks go up	4
round shaped mouth	curved shaped mouth/smile	4
teeth close together	teeth wide apart	3

B. Rating of the Synthetic Characters Animations

In this section we discuss the results of Questionnaire 2. We build our model for the perception of animations of emotions by people using both Principal Component Analysis and Cluster Analysis. Of all the participants 62.5% did not answer the first questionnaire. Our group of participants is aged between 15 and 56 years old, with the mean being

[4]jsPsych version 6.3, a JavaScript library for running behavioral experiments in a web browser, https://www.jspsych.org/6.3/

[5]Google Forms, by Google: https://www.google.com/forms/about/

[6]Unity is a cross-platform game engine developed by Unity Technologies, first announced and released in June 2005. https://unity3d.com

25 years old and a standard deviation of 7.808. Regarding gender 40% are male and 60% are female. As for experience with Synthetic Characters, 80% interacted with Non-Player Characters in videogames and 77.5% interacted with Virtual Agents. As for the interaction with Social Robots only 27.5% and finally of the 40 participants only 2 had worked with or developed Synthetic Characters/Virtual Agents.

1) Principal Component Analysis: We performed Principal Component Analysis for grouping constructs that are correlated. Principal Component Analysis is a variable-reduction technique, it aims to reduce a larger set of variables into a smaller set of variables.

By using SPSS Statistics[7] dimension reduction analyses we run it with a fixed number of factors such as two, three, four and five, in order to find the optimal number of factors. We found that more than five factors did not account for much more variance so we stopped considering further, as for two factors only 57% of variance was not enough so it was also discarded.

We decided on the three factors solution, with 69.9% of the variance. In Table IV we have the rotated component matrix, where two of the constructs cross loaded on more than one component: items with a loading less than 0.4 were removed as recommended and standard process [20]. From this analysis was born the concept of a model which we called *Perception of Animations of Emotions' Model.*

TABLE IV
ROTATED COMPONENT MATRIX FOR 3 COMPONENTS WITH ABSOLUTE VALUE ABOVE 0.4.

Construct		Comp. 1	Comp. 2	Comp. 3
forehead goes down	forehead goes up	0.810		
eyebrows go down	eyebrows go up	0.791		
eyes close	eyes open	0.601		
cheeks move down	cheeks move up		0.815	
mouth remains closed	mouth opens widely	0.420		0.795
round shaped mouth	curved shaped mouth/smile		0.863	
teeth not visible	teeth totally visible			0.907
teeth close together	teeth wide apart	0.526	-0.442	
face contracts	face expands	0.846		

Component 1 reveals that the forehead, eyebrows, eyes, mouth movement, teeth visibility and face usually move in a similar direction, for example something we can clearly see is that the animations for Surprise and Fear would be in one extremity and the animations for Anger and Disgust would be the other. The face expanding is clearly connected with the forehead and eyebrows going up, the eyes and mouth opening and the teeth separating from each other, the characteristics for Surprise and Fear as we will see with more detail in the next section.

Component 2 tells us that the mouth area it is highly correlated with the cheeks and the closeness of the teeth, for example if our SC is smiling, the cheeks would go up and the teeth are closer. Looking at the animation of Happiness we can verify these points.

Component 3 reveals that the more the mouth is opened the more teeth are visible.

To visualize all cases in a 3D space we considered our 3 components model. For each case to be represented we had to find their factor scores. Since a factor is by nature unobserved, we need to first predict or generate plausible factor scores. We used the Regression method in SPSS to generate the factor scores of each case. Figure 3 shows the distribution of the cases labeled with the emotion it represents, each axis represents the different components.

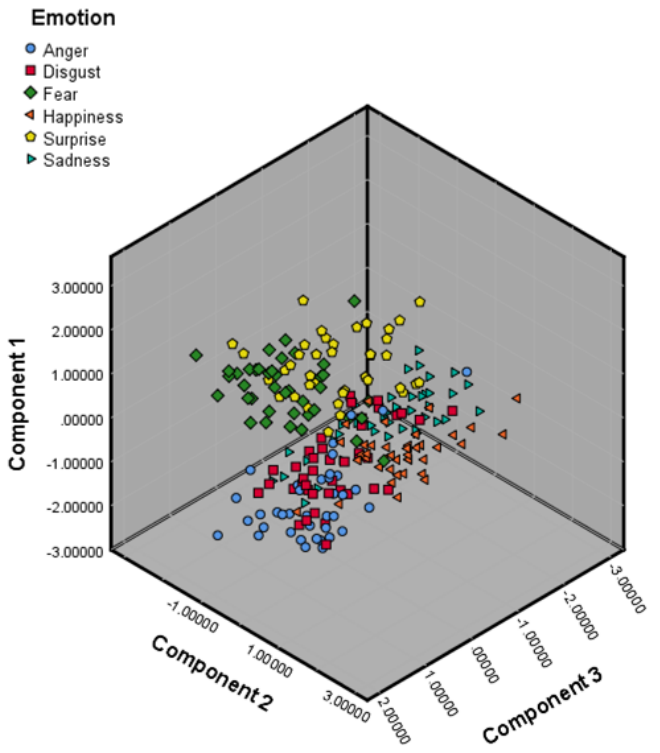

Fig. 3. Distribution of the cases labeled with the emotion it represents, using the calculated factor scores for each case.

2) Cluster Analysis: As already mentioned some emotions are easily confused with one another and with Cluster Analysis we found a way to group them together using our method of the Repertory Grid, we can then identify prototypes of the emotions that are more distinct. Cluster Analysis tries to divide a certain number of observations into different groups that share common characteristics. With a total of 240 cases, with them being the six different animations of emotions multiplied by the number of participants, 40 to be exact, we tried to see

[7]IBM SPSS Statistics version 26 is a software platform that offers advanced statistical analysis. https://www.ibm.com/products/spss-statistics

if different emotions would somehow be put together in the same cluster.

We used SPSS k-means clustering with the squared Euclidean distance method that SPSS provides to minimize within-cluster variances, we considered the factor scores from the 3 components model for the variables and we had a total of 240 cases. Since k-means needs a value for k beforehand we had to understand the optimal number of clusters for our data. We concluded that 4 clusters was ideal for our study by the use of the Elbow and Silhouette Method. Table V, shows the number of cases in each cluster when using the factor scores for each of the 3 components as variables and Table VI the respective cluster centers.

TABLE V
NUMBER OF CASES IN EACH CLUSTER. THE CASES WERE COMPARED WITH THE SQUARED EUCLIDEAN DISTANCE BETWEEN THEIR RESPECTIVE FACTOR SCORES.

Cluster	Anger	Disgust	Fear	Surprise	Sadness	Happiness	Number of cases
1	36	27	0	1	9	1	74
2	0	2	36	29	1	0	68
3	2	3	0	9	27	1	42
4	2	8	4	1	3	38	56
Total	**40**						**240**

TABLE VI
THE CLUSTER CENTERS FOR EACH COMPONENT IN THE PERCEPTION OF ANIMATIONS OF EMOTIONS' MODEL.

	Cluster			
Component	**1**	**2**	**3**	**4**
1	-1.0	1.0	-0.3	0.4
2	-0.1	-0.8	-0.3	1.3
3	0.7	0.3	-1.7	0.0

Cluster 1 is represented by *Anger* and *Disgust* and the main features to consider are: a contraction of the face, in the upper area involving the eyes and forehead (Component 1).

Cluster 2 is represented by *Fear* and *Surprise* and the main features to consider are: the face expanding and the mouth more rounded with the teeth wide apart (Component1 and 2).

Cluster 3 is represented by *Sadness* and the main features to consider are: the mouth more closed with almost no teeth visible (Component 3).

Cluster 4 is represented by *Happiness* and the main feature to consider is the smile (Component 2).

V. CONCLUSIONS

Synthetic Characters are being used more and more everyday, and the recognition of its emotions is a main issue if we want to show more believable characters. The main goal of this study was to understand if we could create a model for how people perceive emotions in specific Synthetic Characters and at the same time try to identify the aspects that may lead to a poor emotion recognition, in particular for non verbal communication such as facial expressions.

We took a set of six animations from specific Synthetic Characters, each representing a different emotion. We made use of a triad comparison to find meaningful constructs, personal rules by which we view or categorize situations, and with these constructs we built a model for how people perceive the different animations.

We first had to understand which constructs were meaningful by requesting to a small group of people to fill out a questionnaire. This questionnaire contained a small section for a demographic characterization of the participants and six sections for the elicitation of the constructs using a triad comparison, a common method for making a Repertory Grid. After analysing a total of 126 constructs using content analysis we reduced it to 9 different constructs.

The second step was to create a model for the perception of animations of emotions, a second questionnaire was necessary. We took the 9 constructs that were selected and made a questionnaire where each animation was to be rated against each construct in a Likert scale with seven points. We used Principal Component Analysis and Cluster Analysis to create the model. Principal Component Analysis reduced the number of constructs to 3 components, the first corresponding to how people viewed the face in general, if it was more contracted or expanded, the second the movement of the mouth and its relation to the cheeks movement and the third and last one the mouth opening and how it is related to the visibility of the teeth. With Cluster Analysis we could see how each different animation was similar as they were grouped together in 4 clusters, Anger and Disgust were grouped in Cluster 1, Fear and Surprise in Cluster 2 and then Sadness and Happiness had their own Cluster, 3 and 4 respectively.

We conclude that with the Repertory Grid method it is possible to see how people perceive different emotions and that it indeed shows us that different emotions are confused with each other due to specific features. We can visualize which features are more prominent in each different cluster and thus how the different animations are more alike. Anger and Disgust are characterized mainly by the first component and the third, Fear and Surprise by the first component and the second, Happiness mainly by only the second and finally, Sadness by the first and the third component.

Furthermore, with the data that we received in our study, it is possible to try to improve emotion recognition of the animations of the Synthetic Characters used in our study. As such the next step would be to make a new set of animations, focusing on the emotions that we saw were similar, such as Anger and Disgust, and Fear and Surprise. We will focus on the constructs that we found, as well as the model that we created using the three components to inform how better distinguish similar emotions (e.g. Anger and Disgust).

A final evaluation has to be made to verify if the new set of animations would really improve emotion recognition. So we would compare the results obtained with an emotion recognition evaluation, one from the base animations used in our study and another with the new animations created. We will verify if emotion recognition was improved. We could also take the changes made by Silva [21] to the animations and compare the three.

Acknowledgment

I would like to thank my supervisor, Prof. Carlos Martinho for the amazing support given throughout the semester, all the insight, expertise and knowledge shared. I would also like to thank Ricardo Rodrigues for all the help and availability.

References

[1] Lisa Feldman Barrett, Ralph Adolphs, Stacy Marsella, Aleix M Martinez, and Seth D Pollak. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological science in the public interest*, 20(1):1–68, 2019.

[2] John N Bassili. Emotion recognition: the role of facial movement and the relative importance of upper and lower areas of the face. *Journal of personality and social psychology*, 37(11):2049, 1979.

[3] Ana Paula Cláudio, Maria Beatriz Carmo, Vítor Pinto, Afonso Cavaco, and Mara Pereira Guerreiro. Virtual humans for training and assessment of self-medication consultation skills in pharmacy students. In *2015 10th International Conference on Computer Science & Education (ICCSE)*, pages 175–180. IEEE, 2015.

[4] Aaron M Curtis, Taylor M Wells, Paul B Lowry, and Trevor Higbee. An overview and tutorial of the repertory grid technique in information systems research. *Communications of the Association for Information Systems*, 23(1):3, 2008.

[5] Artemisa R Dores, Fernando Barbosa, Cristina Queirós, Irene P Carvalho, and Mark D Griffiths. Recognizing emotions through facial expressions: a largescale experimental study. *International journal of environmental research and public health*, 17(20):7420, 2020.

[6] Paul Ekman. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200, 1992.

[7] Paul Ekman. Emotions revealed. *Bmj*, 328(Suppl S5), 2004.

[8] Paul Ekman and Wallace V Friesen. *Unmasking the face: A guide to recognizing emotions from facial clues*. Ishk, 2003.

[9] Valentina Ferretti and Francesco Papaleo. Understanding others: emotion recognition in humans and other animals. *Genes, Brain and Behavior*, 18(1):e12544, 2019.

[10] Fay Fransella, Richard Bell, and Don Bannister. *A manual for repertory grid technique*, chapter 1, 2. John Wiley & Sons, 2004.

[11] Keith Goffin. Repertory grid technique. *Essential Skills for Management Research*, pages 199–225, 01 2002.

[12] Ursula Hess and Pascal Thibault. Darwin and emotion expression. *American Psychologist*, 64(2):120, 2009.

[13] Kyle Johnsen, Robert Dickerson, Andrew Raij, Benjamin Lok, Jonathan Jackson, Min Shin, Jonathan Hernandez, Amy Stevens, and D Scott Lind. Experiences in using immersive virtual characters to educate medical communication skills. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pages 179–186. IEEE, 2005.

[14] Bill Kapralos, Alvaro Uribe-Quevedo, Karen Collins, and Adam Dubrowski. Intelligent avatars and emotion in medical-based virtual learning environments. *Intelligent Decision Technologies*, 13(4):407–416, 2019.

[15] Paul R Kleinginna and Anne M Kleinginna. A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and emotion*, 5(4):345–379, 1981.

[16] Oliver Langner, Ron Dotsch, Gijsbert Bijlstra, Daniel HJ Wigboldus, Skyler T Hawk, and AD Van Knippenberg. Presentation and validation of the radboud faces database. *Cognition and emotion*, 24(8):1377–1388, 2010.

[17] André Lima. Virtual tutor - virtual tutoring agent using empathy and rapport techniques. Master's thesis, Instituto Superior Técnico, University of Lisbon, 2017.

[18] Ricardo Pereira. Expressing emotions through animated speech balloons. Master's thesis, Instituto Superior Técnico, University of Lisbon, 2018.

[19] Ricardo Rodrigues, Ricardo Silva, Ricardo Pereira, and Carlos Martinho. Interactive empathic virtual coaches based on the social regulatory cycle. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 69–75. IEEE, 2019.

[20] P Samuels. Advice on exploratory factor analysis. birmingham city university. *Manuscript in preparation. doi: https://doi. org/10.13140/RG*, 2(5013.9766), 2016.

[21] Ricardo Silva and Carlos Martinho. Supporting affective expression in multiparty interactions. Master's thesis, Instituto Superior Técnico, University of Lisbon, 2018.

[22] Felix B Tan and M Gordon Hunter. The repertory grid technique: A method for the study of cognition in information systems. *MIS quarterly*, pages 39–57, 2002.