An interrelating model between Action Units and Emotional Descriptives: Pleasure, Arousal and Dominance

A lab research project

Daniel Hadar
danielhadar@gmail.com

Guided by Prof. Daphna Weinshall

Hebrew University, Jerusalem

September 9, 2015

1 Project Objective

Developing a model that connects between Action Units (*i.e.* AUs: well known and widely accepted measures of facial expressions) and Descriptive Ratings on three scales: Pleasure, Arousal and Dominance (*i.e.* PAD). Specifically, finding a function from the high-dimensional ($32D$) space of AUs to the 3D-space of PAD ($f : AU \rightarrow PAD$). Ideally this function would be broaden into a tool that receives a video clip of some facial expression(s) as input, and outputs a 3D signal over time in PAD coordinates. Such a tool is meaningful to the field of psychiatry (and mental disorders diagnostic in particular), since non-verbal behavior (and more specifically, facial expressions) are of high importance to psychiatric diagnoses of mental disorders, since they provide diagnostic-related information about the patient that isn't subjected to subjectiveness of the patient's self-report (Philippot, Feldman & Coats, 2003).

2 Introduction

2.1 Classification of Emotion and Facial Expressions

Human facial expressions are a fundamental aspect of our every-day life, as well as a popular research field for many decades. It is not always thought-about or spoken-of, but many of the decisions we take are based upon other people's facial expressions - for example, deciding whether to ask someone's phone number ("is he smiling back?"), or whether a person is to be trusted. Psychologists have been claiming assiduously that human beings are in fact *experts* of facial expressions (as a car-enthusiast would be an expert of car models), in terms of being able to distinguish between highly similar expressions, to recognize large amount of different faces (of friends, family, co-workers, etc.) and being sensitive to subtle facial gestures - because ever since one is born he in fact refines these abilities - he encounters numerous faces throughout he's life, and is required to derive such information from them on a daily basis; That theory had influenced Nueroscientists to locate and specify an area in the brain that's in charge of the perception of faces - the *fusiform area* (Kanwisher, McDermott & Chun, 1997 and Kanwisher & Yovel, 2006).
Since over a 100 years ago, scholars have been interested in finding a rigorous model for the classification of emotions, *i.e.* being able to express and explain every emotion by a closed and minimal set of *primary elements*. The work in this field could be divided into two approaches: **The Categorical Approach** and **The Dimensional Approach** (a comprehensive survey could be found in: Gunes, Schuller, Pantic & Cowie, 2011). To be noted that the approaches are non-contradictory, but merely propose a different description methods.

2.1.1 The Categorical Approach

The general idea that underlies in the basis of this approach is that there's a set of *Basic Emotions* that are innate, and every other emotion could be described as a combination of them. Many scholars that support this theory would add another claim - that these emotions are *universal*, *i.e.* they are (almost) identical across cultures.
The Categorical approach's foundations were set back in 1872 by Charles Darwin, who proposed that

universal facial expressions exist, and are inherited. Darwin claimed that until a certain point in time facial expressions had some biological origin, and throughout the generations the expressions became innate (Darwin, 1872). Over time came many scholars that fortified and modified this theory: In 1924 Floyd Allport claimed that beyond the existence of *universals*, the biological mechanism in fact still exist (*e.g.* twitching the face in the presence of bad odor), saying: *"Instead of the biologically useful reaction being present in the ancestor and the expressive vestige in the descendent, we regard both this functions as present in the descendent..."* (Allport, 1924). In 1971 Paul Ekman gathered the scattered ideas about universal facial expressions and combined them into a unified theory that includes a list of these basic emotions: Surprise, Anger, Disgust, Fear, Interest, Sadness and Happiness (*i.e.* *SAD FiSH*), and supported it with several cross-cultural experiments (Ekman, 1971). Over the years Ekman and his colleagues refined this theory, by carrying out more experiments (some even included the use of EMG measures to show correlation between mood and facial muscle activity (Ekman, 1992a)) and changing bits and pieces in it, as eventually *Interest* was omitted (Ekman, 1992b).

2.1.2 The Dimensional Approach

This approach differs from The Categorical Approach by the *primary elements* that compose it - while in the Categorical Approach the elements are basic prototypical emotions (*SAD FiSH*), in the Dimensional Approach the elements are *bipolar-diagonal Axes* (or *dimensions*) that every emotion could be described as a combination of them (and could be thought of as a Basis of a Vector space in Linear Algebra).

The Dimensional Approach was first proposed in 1897 by who is considered to be "the father of modern psychology" (Carlson, Heth, Miller, Donahoe, & Martin, 2009) Wilhelm Maximilian Wundt, that suggested that emotions could be described by three dimensions: *"Pleasurable-Unpleasurable"*, *"Arousing-Subduing"* and *"Strain-Relaxation"* (Wundt & Titchener, 1897). In 1941 Harold Schlosberg suggested his similar version of *continuous* description of emotions and *facial expressions* (Schlosberg, 1941), and 10 years later (after conducting several experiments to reinforce it), released it as a unified theory of two-dimensions: *"Pleasantness-Unpleasantness"* and *"Attention-Rejection"* (Schlosberg, 1952), which was later extended to three dimensions, adding *"level of activation"* (as a "possible third dimension") (Schlosberg, 1954).

Similarly to the Categorical Approach, some modifications were also suggested for the Dimensional Approach over the years, as several scholars proposed some change of the axes (*e.g.* adding *"Sleepy-Aroused"* or *"Dominant-Submissive"* (Russell & Mahrabian, 1977), adding a forth axis: *"Exciting-Glooming"*, *"Distressing-Relaxing"* or even *"Positive-Negative"* (Russell, 1980 and Remington, Fabrigar & Visser, 2000), combining the Categorical models with the Dimensional model and creating Hybrid Models (Plutchik, 1980, Plutchik 2001 and Panayiotou, 2008), or simply conducting more experiments to fortify it and reemphasize the *diagonality* of the axes (Abelson & Sermat, 1962 and Fontaine, Scherer, Roesch, & Ellsworth, 2007); And after examining dozens of articles, we can say with confidence that there's a general consensus about the three following dimensions: *"Pleasure-Displeasure"*, *"Arousal-Nonarousal"* and *"Dominance-Submissiveness"*.

2.2 Facial Action Coding System (FACS)

In order to address facial expressions with mathematical tools, one must first be able to define them using some measurable description system. In 1976 Ekman & Friesen introduced a system called Facial Action Coding System ($FACS$) that was developed to measure and describe facial movements (Ekman & Friesen, 1976). The system's basic components are primary elements called **Action Units** (AUs), where most AUs are associated with a facial muscles (*e.g.* $AU-2$ with *Outer Brow Raiser*) and some describe more general gestures (*e.g.* $AU-19$ means *Tongue Out*). Every facial expression could be described by the set of AUs that compose it (*e.g.* $Fear \equiv \{1, 2, 4, 5, 7, 20, 26\}$). Over the years Ekman and his colleagues expended the set of AUs (and their expressiveness accordingly), adding head movements, eye movements, gross behavior (*e.g.* *Sniff* and *Head-Nod*), etc. Currently there are 86 AUs, 56 of them are solely face-related.

Describing a facial expression in terms of FACS is traditionally done by FACS-experts; Becoming one requires either taking a course or studying the subject independently from guides and book provided by Ekman's Lab[1]. In the recent years, the growth of *Computer Vision* yielded several **automatic** tools for facial expressions recognition (*i.e.* *Affect Recognition*); An exhaustive but comprehensive survey (that includes recognition of audio as well) could be found in: Zeng, Pantic, Roisman & Huang, 2009.

3 Current Project

As mentioned, the goal of this project was to find a model M that maps values from the AU space to the PAD space (for example (fake): $M(2, 4, 14, 34, 35) = (1, 4, 7)$, where the output is interpreted as: Pleasure = 1, Arousal = 4, Dominance = 7). In other words, the goal was to find a function $M : AU \to PAD$ that reduces the data's dimension ($\mathbb{R}^{32} \to \mathbb{R}^3$) in order to use FACS data in a more descriptive manner, in terms of behavioral psychology. To be noted that machine learning dimension reduction tools (like PCA) are shortcoming for this goal; Thus one can look at this mission as an attempt for some "smart" dimension reduction that's done in order to gain data in lower dimension that has a real-life meaning.

The easiest way would be finding a dataset of images or clips that includes both labellings (AUs and PAD ratings) for each image/clip. To our disappointment, such a dataset was nowhere to be found - so we were compelled to create one of our own. To do so we used an images-sequence dataset of posed facial expressions called Cohn-Kanade (CK) (Kanade, Cohn & Tian, 2000 and Lucey, Cohn, Kanade, Saragih, Ambadar & Matthews, 2010) in which every image-sequence comes with its corresponding AUs set that's present in it. We used the public-domain Java image processing program ImageJ[2] to turn each images-sequence to a short video clip (~ 3 seconds). Note that in order to use a dataset that contains PAD ratings and achieve FACS ratings, we would need to use services FACS experts (which are rare and expensive); In contrast, to achieve PAD ratings we only needed to get a number of subjects large enough (in statistical manner) and use their mean ratings.

[1] http://www.paulekman.com/products/
[2] http://imagej.nih.gov/ij/index.html

3.1 Experiment

3.1.1 Subjects

20 Students from Jerusalem (7 males and 13 females), ages: 15-29, during the first half of 2015. Each student was given the same instructions (*see 3.1.3*) and was paid 30 NIS. The experiment was carried out in Givat Ram Campus of The Hebrew University in Jerusalem.

3.1.2 Procedure

Each subject was presented with 175 CK clips (35 clips in 5 sets) and was asked to rate them over 3 dimensions (Pleasure, Arousal and Dominance), each in scale of 1-9. The display of the clips and their ratings was done using a MatLab program[3].

3.1.3 Instructions

At the beginning, each subject received a paper of instructions in English; They could be found in the *Appendix*.

3.2 Model Construction

As mentioned, the original goal was to build a model from AU space to PAD space. Eventually we had build **2 models** - one as described ($AU \rightarrow PAD$) and another one in the opposite direction ($PAD \rightarrow AU$). The latter model provides us with AUs that are better predicted from PAD ratings, and thus could be used for example for feature selection for the former model.

3.2.1 $PAD \rightarrow AU$

The model is a function $M : \left(\{[1,9]\}^3 \times AU_{\in[1,64]}\right) \rightarrow \{\pm 1\}$ defined as follows: given some PAD ratings (*i.e.* a tuple of numbers $(P,A,D)_{\{P,A,D\}\in[1,9]^3}$) and an AU (*i.e.* a number $a \in [1,64]$), the model returns whether a is (or more precisely - *is expected to be*) present in the facial expression that was rated (P,A,D). In order to do so, for each AU we used L2-regularized L2-loss SVC:

$$min_w \frac{w^T w}{2} + C \sum \left(max(0, 1 - y_i w^T x_i\right)^2$$

Carried out in MatLab, using a designated package called LIBLINEAR (Fan, Chang, Hsieh, Wang and Lin, 2008)[4] to learn the data; Thus we had built 32 models (for each AU): $M_{AU}\left(P,A,D\right) \rightarrow \{\pm 1\}$.

3.2.2 $AU \rightarrow PAD$

Definition: A *Binary Vector of AUs* is a 32 entries binary vector that defines the AUs present in a facial expression; For example, for a clip with the following AUs: 4,7,10,17,23,24,43, the corresponding

[3] https://github.com/danielhadar/vid_project
[4] Software available at: http://www.csie.ntu.edu.tw/∼cjlin/liblinear

Binary Vector of AUs is:

$$[0,0,1,0,0,1,0,1,0,0,0,0,0,0,1,0,0,0,0,1,1,0,0,0,0,0,0,0,1,0,0,0]$$

Notice that although the minimal AU is 1 and the maximal is 64, there are only 32 AUs:

$$[1,2,4,5,6,7,9,10,11,12,13,14,15,16,17,18,20,21,22,23,24,25,26,27,29,30,38,39,43,45,54,64]$$

The model is a function $M:\mathbb{R}^{32}\to\mathbb{R}^3$ defined as follows: given some *Binary Vector of AUs*, return the PAD rating (*i.e.* the tuple $(pleasure, arousal, dominance)_{(p,a,d)\in[1,9]^3}$) of its corresponding facial expression, based on a linear model achieved by linear regression w.r.t square-loss error:

$$min_w \frac{w^T w}{2} + C\sum max(0, |y_i - w^T x_i|)^2$$

We aimed to build a model in 2 different ways: one was to build the model by connecting 3 axis-specific models (*i.e.* build $M_s:\mathbb{R}^{32}\to\mathbb{R}$ for each $s\in\{pleasure, arousal, dominance\}$ and compose: $M(v) = (M_{pleasure}(v), M_{arousal}(v), M_{dominance}(v))$), and the other was by connecting one axis-specific model with a model for 2-axis computed together, using multivariate regression (detailed under *results*). To be stated that the former yielded better results.

4 Results

4.1 Rating's Average, Median and Internal Consistency

Using *Cronbach's alpha* (Santos, 1999 and Bland & Altman, 1997) we measured the internal consistency (*i.e.* average of correlations) of the clips ratings (each dimension separately) over the subjects. The results, along side the averages and medians over the entire dataset could be seen in *Table 1* (Note that internal consistency ≥ 0.9 is considered excellent).

	α	Average	Median
Pleasure	0.953	4.00	3.77
Arousal	0.955	5.22	5.6
Dominance	0.950	4.46	4.47

Table 1: Rating's Average and Median

4.2 Models

4.2.1 *PAD → AU*

As said, a model was computed for each AU separately (32 models in total). Each model was evaluated using *Leave-One-Out* Cross-validation (*i.e.* 20 runs over 19 subjects $\left\{1...\hat{i}...20\right\}$ $\forall i\in[20]$ and prediction over the subject i that was left out). Since the data in unbalanced (*i.e.* some AUs appear significantly more than others) we used *Matthews correlation coefficient* (Matthews, 1975) to evaluate the models accuracy:

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)\,(TP+FN)\,(TN+FP)\,(TN+FN)}}$$

The average MCC grade across all the models is $\overline{MCC} = 0.141$, $\sigma_{mcc} = 0.16$, withing the range $[-0.0021, 0.5388]$. The results show that some AUs are more predictable than others: on one hand, the highest 20%'s MCC grade is over 0.33, while on the other hand there are 14 AUs ($\sim$ 44% of all AUs) that their MCC is zero or less (*i.e.* are not better than random prediction). The full results could be found in the *Appendix.*

4.2.2 $AU \rightarrow PAD$

As said, a model was computed for each axis separately (pleasure, arousal, dominance). Each model was evaluated using *Leave-One-Out* Cross-validation (*i.e.* 175 iterations of leaving one clip out in each iteration), and the average and standard deviation of its mean square error were calculated, as well as the average correlation coefficient (Pearson's r) between the predicted values and the actual values. The resutls could be seen in *Table 2.*

	Average	STD	ρ
Pleasure	1.306	2.714	0.7349
Arousal	1.591	2.365	0.7634
Dominance	1.341	2.171	0.6019

Table 2: Average, Standard Deviation and Prearson's r for the model's predictions

Next, we attempted using multivariate regression for pairs of axes with high correlation. The only pair with a reasonably good correlation was *arousal-dominance* ($\rho_{pleasure,arousal} = 0.233$, $\rho_{pleasure,dominance} = 0.1573$, $\rho_{arousal,dominance} = 0.7862$). A model was constructed for it, but yielded worse results ($r_{arousal} = 0.588$, $r_{dominance} = 0.375$).

5 Discussion

Interestingly, there's no obvious mutual *physical* line between all high-rated AUs (*i.e.* most predictable), but they vary between almost every facial area: chin, mouth, philtrum (mustache area), eyebrows, lids and lips. That being said, it is worth noting that they could be grouped as the more *significant* uses of these muscles, and that there's a high correlation between the expressions derived from their anatomical definition and the ratings they predict:

- **Lip Corner Puller** (Zygomaticus Major, $AU-12$, $MCC = 0.5388$). Pulls the angles of the mouth upwards, outwards and backwards and deepens the wrinkles under the eyes; Involved in expressions of joy and happiness: smile and laughter (Artist, 1991) (*See Figure 1*). Correspondingly, $AU-12$ is a great predictor for ***high-pleasure*** expressions - the average pleasure rating for clips that it's present in is high by 2.26 from the total average, while the average for arousal and dominance is almost identical to their average. Moreover, in the clips in which it isn't present, the average for arousal and dominance isn't hurt while the average for pleasure decreases by 0.45, as could be seen in *Table 3*.

	μ^{total}	μ^{AU-12}	$diff.$	$\mu^{!AU-12}$	$diff.$
Pleasure	4.00	6.26	+**2.26**	3.55	**-0.45**
Arousal	5.22	5.20	**-0.02**	5.24	+**0.02**
Dominance	4.46	4.60	+**0.14**	4.42	**-0.04**

Table 3: Averages for all clips (μ^{total}), clips with $AU-12$ (μ^{AU-12}) and clips without $AU-12$ ($\mu^{!AU-12}$)

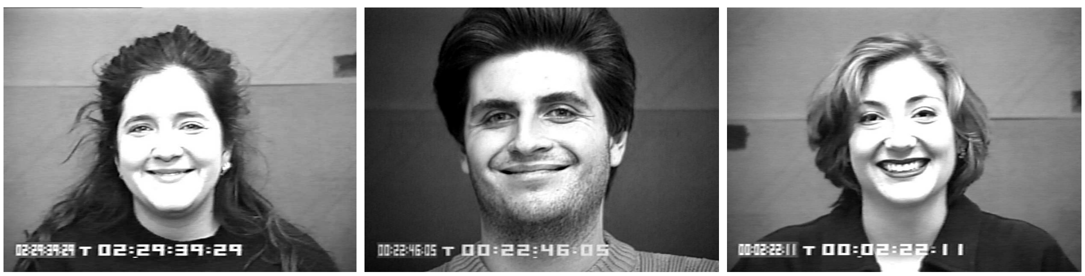

Figure 1: Expressions with high presence of $AU - 12$ (Lip Cornet Puller)

- **Chin Raiser** (Mentalis, $AU - 17$, $MCC = 0.3976$). Raises the fleshy prominence of the chin, that pushes the lower lip upwards; Involved in expressions of sadness, grief, anger and disdain, and is prominent when one tries to conceal extreme sadness (crying) by forcing the mouth to stay closed (*ibid*). (*See Figure 2*). Correspondingly, $AU - 17$ is a good predictor for ***low-pleasure, low-arousal*** expressions - their average when it is present is lower by 1.08 and 0.9 respectively from the average; and when they're not, it's higher by 0.55 and 0.49, as could be seen in *Table 4*. It's worth noting that posing an expression of real restrained sadness or anger is not an easy assignment, thus many of them are seemingly fake. Moreover, our dataset contains a small amount of clips with such expressions - thus it seems reasonable to assume that over a real-life dataset (non-posed expressions and more of these type of clips) the quality of this predictor would be even more significant.

	μ^{total}	μ^{AU-17}	$diff.$	$\mu^{!AU-17}$	$diff.$
Pleasure	4.00	2.92	**-1.08**	4.55	+**0.55**
Arousal	5.22	4.32	**-0.9**	5.71	+**0.49**
Dominance	4.46	4.21	**-0.25**	4.58	+**0.12**

Table 4: Averages for all clips, clips with $AU - 17$ and clips without $AU - 17$

Figure 2: Expressions with high presence of $AU-17$ (Chin Raiser)

- **Upper Lid Raiser** (Levator Palpebrae Superioris, $AU-5,\ MCC=0.3479$). Raises the upper eyelid; Tends to produce ***intense*** expressions, ranging from fear and shock to intense joy or a blank gazing expression (*ibid*). (*See Figure 3*). Not surprisingly, $AU-5$ is a great predictor for intense expressions that reflect in ***high-arousal***, while the pleasure and dominance are closer to their averages, but still are higher when it is present; Moreover, it's absence affects ***arousal*** mostly as would be expected from expressions with partially or fully closed eyes, as could be seen in *Table 5*.

	μ^{total}	μ^{AU-5}	$diff.$	$\mu^{!AU-5}$	$diff.$
Pleasure	4.00	4.43	+**0.43**	3.83	**-0.17**
Arousal	5.22	6.97	+**1.75**	4.61	**-0.61**
Dominance	4.46	5.09	+**0.63**	4.22	**-0.24**

Table 5: Averages for all clips, clips with $AU-5$ and clips without $AU-5$

Figure 3: Expressions with high presence of $AU-5$ (Upper Lid Raiser)

- **Brow Lowerer** (Corrugator Supercilii, $AU-4,\ MCC=0.3413$). Pulls the middle section of the eyebrows downward and medially in an oblique direction; Produces expressions of suffering and pain on one hand, and concentration and confusion on the other (*ibid*). (*See Figure 4*). Correspondingly, $AU-4$ is a great predictor for ***low-pleasure*** (1.19 below the average), while the arousal and dominance are almost identical to their averages, as could be seen in *Table 6*.

	μ^{total}	μ^{AU-4}	$diff.$	$\mu^{!AU-4}$	$diff.$
Pleasure	4.00	2.81	**-1.19**	4.44	+**0.44**
Arousal	5.22	5.22	**0**	5.24	+**0.02**
Dominance	4.46	4.74	+**0.28**	4.34	**-0.12**

Table 6: Averages for all clips, clips with $AU-4$ and clips without $AU-4$

Figure 4: Expressions with high presence of $AU-4$ (Brow Lowerer)

Secondly, it's possible that there's a more informative way to describe the AUs than binary vectors; Thus there might be some better models based on the data collected. Further research can use the gathered data and apply some other forms of learning. Another further research that could be carried out is to use non-posed clips. When we started this project there where no such available that suited our needs; Now there are, released by the same lab (Cohn-Kanade (CK)). Non-posed clips would probably yield results that would be more similar to real-life observations over facial expressions; For example, $AU-26$ (Jaw Drop) would tend to appear in high-arousal expressions, mostly in ones that involve surprise (Russell & Barrett, 1999), but in our data they are surprisingly (pun intended) associated with low-arousal expressions ($\mu^{AU-26}_{arousal} = 4.6$, 0.62 below average, while $\mu^{AU-26}_{pleasure} = 4.0$ and $\mu^{AU-26}_{dominance}4.49$ are almost identical to the average). This could be explained by watching the clips in which this AU is present - because the expressions are posed by non-actors, the drop of the jaw seems clearly forced, which gives it an unnatural feeling and reflect a more despising or snobbish expression than a surprised one (*see Figure 5*).

Figure 5: Expressions with high presence of $AU-26$ (Jaw Drop)

Lastly, it's worth noting that there's a high correlation between the number of appearances of an AU in the dataset and it's MCC grade (*i.e.* how good of a predictor is it), As could be seen in *Figure*

6 ($\rho_{\#,MCC} = 0.76$, the green line). This observation makes sense because MCC grade is intolerant for all-right/all-wrong predictions (the nominator becomes zero), and rare AUs tend to produce such models (training over a small amount of examples). Further research could be done over an even more uniformly distributed data in terms of AUs.

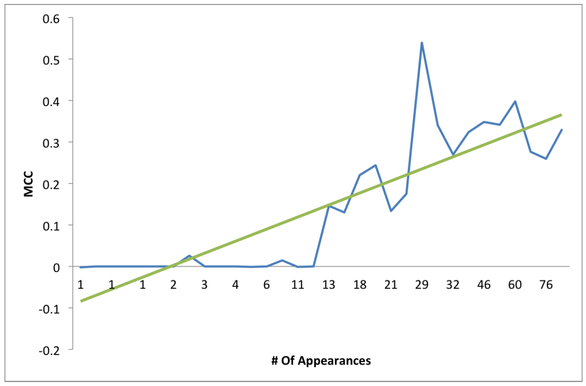

Figure 6: MCC grade as a function of Number Of Appearances of an AU

6 Bibliography

1. Abelson, R. P., & Sermat, V. (1962). Multidimensional scaling of facial expressions. Journal of Experimental Psychology, 63(6), 546.

2. Allport, F. H. (1924). Scientific Books: Social Psychology. Science, 59, 440-441.

3. Artist, E. G. (1991). Human Anatomy for Artists: The Elements of Form: The Elements of Form. Oxford University Press.

4. Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. Bmj, 314(7080), 572.

5. Carlson, N. R., Heth, D., Miller, H., Donahoe, J., & Martin, G. N. (2009). Psychology: the science of behavior. Pearson. p. 18.

6. Darwin, C. (1872). The expression of the emotions in man and animals. London: John Murray.

7. Ekman, P. (1971). Universals and cultural differences in facial expressions of emotion. Nebraska Symposium on Motivation, Vol 19, 1971, 207-283.

8. Ekman, P. (1992a). An argument for basic emotions. Cognition & emotion, 6(3-4), 169-200.

9. Ekman, P. (1992b). Facial expressions of emotion: New findings, new questions. Psychological science, 3(1), 34-38.

10. Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. Environmental psychology and nonverbal behavior, 1(1), 56-75.

11. Fan, R. E., Chang, K. W., Hsieh, C. J., Wang, X. R., & Lin, C. J. (2008). LIBLINEAR: A library for large linear classification. The Journal of Machine Learning Research, 9, 1871-1874.

12. Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. Psychological science, 18(12), 1050-1057.

13. Gunes, H., Schuller, B., Pantic, M., & Cowie, R. (2011, March). Emotion representation, analysis and synthesis in continuous space: A survey. In Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on (pp. 827-834). IEEE.

14. Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, 46-53.

15. Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. The Journal of Neuroscience, 17(11), 4302-4311.

16. Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. Philosophical Transactions of the Royal Society B: Biological Sciences, 361(1476), 2109-2128.

17. Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101.

18. Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta (BBA)-Protein Structure, 405(2), 442-451.

19. Panayiotou, G. (2008). Emotional dimensions reflected in ratings of affective scripts. Personality and Individual Differences, 44(8), 1795-1806.

20. Philippot, P., Feldman, R. S., & Coats, E. J. (Eds.). (2003). Nonverbal behavior in clinical settings. Oxford University Press.

21. Plutchik, R. (1980). Emotion: A psychoevolutionary synthesis. Harpercollins College Division.

22. Plutchik, R. (2001). The Nature of Emotions Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. American Scientist, 89(4), 344-350.

23. Russell, J. A. (1980). A circumplex model of affect. Journal of personality and social psychology, 39(6), 1161.

24. Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. Journal of personality and social psychology, 76(5), 805.

25. Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. Journal of research in Personality, 11(3), 273-294.

26. Remington, N. A., Fabrigar, L. R., & Visser, P. S. (2000). Reexamining the circumplex model of affect. Journal of personality and social psychology, 79(2), 286.

27. Santos, J. R. A. (1999). Cronbach's alpha: A tool for assessing the reliability of scales. Journal of extension, 37(2), 1-5.

28. Schwartz, G. E., Fair, P. L., Salt, P., Mandel, M. R., & Klerman, G. L. (1976). Facial muscle patterning to affective imagery in depressed and nondepressed subjects. Science, 192(4238), 489-491.

29. Scholsberg, H. (1941). A scale for the judgment of facial expressions. Journal of experimental psychology, 29(6), 497.

30. Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. Journal of experimental psychology, 44(4), 229.

31. Schlosberg, H. (1954). Three dimensions of emotion. Psychological review, 61(2), 81.

32. Wundt, W., & Titchener, E. B. (1897). Scientific Literature: Outlines of Psychology; An Outline of Psychology. Science, 5, 882-884.

33. Zeng, Z., Pantic, M., Roisman, G., & Huang, T. S. (2009). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 31(1), 39-58.

7 Appendix

7.1 MCC grades and number of appearances, for each AU

AU	MCC	# Of Appearances
18	-0.0021	1
45	-0.0017	6
11	-0.001	11
10	-0.0005	3
16	0	6
21	0	1
22	0	1
26	0	13
29	0	2
30	0	1
39	0	4
43	0	3
54	0	1
64	0	1
38	0.0145	7
13	0.0263	2
14	0.1304	17
20	0.1337	21
9	0.1462	13
6	0.1757	25
24	0.22	18
23	0.2443	19
1	0.2594	76
7	0.2695	32
2	0.2767	67
27	0.3231	41
25	0.3291	100
15	0.34	30
4	0.3413	48
5	0.3479	46
17	0.3976	60
12	0.5388	29

Table 7: MCC grades and number of appearances for each AU

7.2 Instructions

In the following experiment you will watch several clips of people posing facial expressions to the camera.

You will be asked to rate them on 3 scales (1-9):

(1) Pleasure-Displeasure.

(2) Arousal-Nonarousal.

(3) Dominance-Submissiveness.

A Few Definitions:

1. Pleasure-Displeasure = How pleasant is the emotion? For Example: both anger and fear are unpleasant emotions, while joy is a pleasant emotion.

[Here came 2 example of high-pleasure and 2 examples of high-displeasure]

2. Arousal-Nonarousal = How intense is the emotion? For Example: while both anger and rage are unpleasant emotions, rage is more intense than anger; On the other hand, boredom has a low arousal value (although it is also unpleasant).

[Here came 2 example of high-arousal and 2 examples of high-nonarousal]

3. Dominance-Submissiveness = How controlling and approaching the emotion is? For Example: while both anger and fear are unpleasant emotions, anger is dominant but fear is submissive.

[Here came 2 example of high-dominance and 2 examples of high-submissiveness]