Refactoring Facial Expressions: an Automatic Analysis of Natural Occurring Facial Expressions in Iterative Social Dilemma

Giota Stratou*, Job Van Der Schalk†, Rens Hoegen* and Jonathan Gratch*

*University of Southern California, Institute for Creative Technologies, Los Angeles, California
Email: {stratou, rhoegen, gratch }@ict.usc.edu
†Cardiff University, Wales, UK
Email: vanderSchalkJ@cardiff.ac.uk

Abstract—Many automatic facial expression recognizers now output individual facial action units (AUs), but several lines of evidence suggest that it is the combination of AUs that is psychologically meaningful: e.g., (a) constraints arising from facial morphology, (b) prior published evidence, (c) claims arising from basic emotion theory. We performed factor analysis on a large data set and recovered factors that have been discussed in the literature as psychologically meaningful. Further we show that some of these factors have external validity in that they predict participant behaviors in an iterated prisoner's dilemma task and in fact with more precision than the individual AUs. These results both reinforce the validity of automatic recognition (as these factors would be expected from accurate AU detection) and suggest the benefits of using such factors for understanding these facial expressions as social signals.

1. Introduction

Researchers in affective computing face a dilemma when choosing or developing methods to classify facial expressions. Echoing controversies within emotion psychology, some techniques categorize the face in terms of basic emotions such as joy, fear and anger [1], whereas others identify core affect such as valence and arousal [2]. Complicating this distinction, some expressions may have nothing to do with emotion at all [3], yet serve important communicative functions that affectively intelligent systems may still wish to identify. As a result, modern emotion recognition techniques are increasingly focusing on more primitive face representations from which higher level abstractions, such as basic emotions, can be constructed. The most prominent of these approaches is to recognize individual facial action units [4], or AUs, such as the Lip Corner Puller (AU12) or Brow Lowerer (AU4). Many research and commercial expression recognition techniques now classify the face in terms of these primitive actions, as well as some higher level representations.

In this paper, we investigate an approach that falls between the extremes of low-level action units and high-level basic expressions. We perform a factor analysis of individual action units over several very large corpora of facial expressions. The result yields two interesting contributions for affective computing. First, we show that action units

cluster into a small number of meaningful factors that have been previously discussed in the psychological literature (based on analysis performed with expert hand annotations). This provides evidence for concurrent validity in that combinations of automatically recognized action units represent socially and psychologically meaningful expressive displays, as opposed to being artifacts of the machine learning approach used to identify them. Second, we show that these factors have external validity in that they are more predictive of a person's behavior than individual action units alone. Specifically, we show that some of the identified factors are predictive of a person's actions in an iterated prisoner's dilemma game.

In Section 2, we discuss current views on facial expression classification and work on automatic analysis in social dilemmas. In Section 3, we present the factor analysis from large data. In Section 4, we describe the corpus we used for analysis, and the results from correlating behaviors with facial expressions in this context. We include further discussion in Section 5 and close with conclusions.

2. Facial Expressions in Naturalistic Data

In his influential 1872 book, Charles Darwin introduced the argument that facial expressions are evolved and adaptive [5]. They not only evolved as part of internal emotional systems to prepare an organism to act (as widening the eyes brings in more information when the animal is surprised) but have important communicative functions (as displays of anger and supplication can deflect costly conflicts). Since Darwin, emotion researchers have fiercely debated how expressions look, what they signify, and how they shape social interaction. Affective computing doesn't settle these debates but brings the opportunity to reexamine them in the context of massive quantities of data.

2.1. What do Expressions Signify?

When analyzing a face, automatic facial expression techniques must generate one or more labels that characterize the facial pose. Thus, immediately, affective computing must adopt a theoretical perspective on what this facial pose signifies. Affective computing research has been heavily influenced by Ekman's notion of six basic emotions: anger, disgust, fear, joy, sadness, surprise [6]. Initial evidence for

these expressions emphasized that cultures across the world could universally recognize posed facial displays of these expressions, and early work in affective computing showed that automated techniques were also quite accurate in recognized posed or acted displays of these basic expressions.

Unfortunately, facial expressions in natural interactions are far more complex. People rarely show static or complete basic expressions and the effectiveness of these labels is questionable [7], [8], [9]. To study this complexity, Ekman and Friesen developed the Facial Action unit Coding System (FACS) [4], which offers higher resolution and yields mainly morphological information, leaving room for interpretation based on context. Facial action units (AUs) better capture the subtlety of expressions in natural interactions, although they defer the question of what these expressions signify. For example, AU1(inner eyebrows up) could be part of surprise but also signify fear. Ekman's hope was to show that evidence of basic emotions would emerge from an analysis of large quantities of naturalistic data.

The evidence for Ekman's claims is mixed and open to conflicting interpretations. It is now clear that full-blown prototypical displays of emotion rarely occur in natural interactions. Some interpret these results as evidence that people are masking or regulating their authentic expressions. For example, Keltner introduced the notion of smile controls (meaning a combination of action units that mask smiles) as evidence that people were regulating their true feelings [10]. Others such as Barrett and Russell viewed the evidence as undermining the validity of basic expressions and argued that expressions merely signify valence and arousal, or what they call core affect [11].

Affective computing, as a whole, has been relatively agnostic to this controversy. Some expression recognizers output the six basic emotions, others output valence, yet others offer a combination of outputs. Increasingly, trackers such as CERT [12] and Openface [13] offer facial action units as well as these more abstract representations, allowing developers to make use of whatever works for their application. However, the very success of automatic recognition can allow us to revisit Ekman's original question. In the current paper, we consider a bottom-up approach, using factor analysis, to see if coherent patterns of facial displays begin to emerge when analyzing AUs at a large scale in naturalistic data. Data mining approaches to analyzing spontaneous facial expressions have been used in previous studies and it has indeed been demonstrated that combinations of AUs can account for more variation in behavior than AUs alone [14]. However, to our knowledge our work is the first systematic attempt to recover basic expressions from a large dataset of naturally occurring facial behavior.

2.2. What do Expressions do?

One way to assess the validity of alternative representations of expressions is to shift the focus. Rather than focusing on what expressions signify, an alternative approach is to consider what they do. As Darwin claimed, expressions may have evolved and hold adaptive value in shaping the nature of social interactions. Regardless of their

TABLE 1. DATASETS USED FOR EXPRESSION FACTOR ANALYSIS

Dataset	#Sessions	#Frames	Description
IPD	608	1.3m	Iterated social dilemma (current analysis data) [22]
DAIC-F2F	160	5.4m	Distress assessment interviews [23]
CRA-F2F	364	2.9m	Dyadic negotiation task [24]

connection to underlying emotional states, basic expressions of emotion can have value to affective computing if they reliably predict outcomes in social interaction. Following this logic, alternative representations can be compared in terms of how much variance they explain in social behavior.

The social function of emotions has been most heavily studied in the context of standard game-theoretic tasks such as social dilemmas and negotiations [15], [16], [17], [18]. Some of the most robust findings come from the field of psychology and cross-disciplinary studies and mention that anger and joy are prominent in negotiations. Also the lack of expression reciprocity has been found important for dyadic cooperation [17].

The iterated prisoners' dilemma (IPD) [19] provides an interesting opportunity to study the social functions of facial expressions because it creates a dilemma between cooperation and competition. Partners can do well if they both collaborate but the incentive structure of the task creates an incentive to exploit one's partner and people may choose non-cooperation out of greed, but also through fear of exploitation. The game-theoretic solution is to always choose non-cooperation (it is the only Nash equilibrium [20]) yet human participants show high-levels of cooperation. Researchers have argued that emotional expressions allow people to solve this dilemma by communicating important information about the partner's emotions and intentions [21].

Within our current work, we investigate facial behaviors across different choices and outcomes in a finite horizon iterated social dilemma task. Specifically we look at automatically extracted facial behaviors that are more refined than the basic emotions. We perform analysis with the most refined unit of automatic analysis (facial action units) and create new constructs (factors) based on general context. We then evaluate the effectiveness of those factors in this context.

3. Factors

We aim to create new constructs for facial expression analysis based on facial action unit output from automatic trackers. For this purpose, we apply factor analysis on a large set of video data ($\sim$10m frames, sampled at 0.06s) of naturalistic expressions. We investigate the validity of those factors based on i) their mapping into meaningful constructs from literature and ii) their external validity, by correlating with meaningful phenomena in our data.

We use 3 large datasets (seen in Table 1). The *IPD* data consists of 608 participants in an iterated prisoner's dilemma over a webcam mediated setup (described in more detail in the next section). In the *DAIC* dataset, participants

TABLE 2. Facial Expression Factors. Loading Cutoff=0.3

|	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
AU01		0.981				
AU02		0.873				
AU04					0.703	
AU05						
AU06	0.982					
AU07	0.466				0.384	
AU09					0.390	
AU10						0.488
AU12	0.882					
AU14				0.724		
AU15						0.504
AU17				0.632		0.420
AU20			0.377			
AU23				0.710		
AU25			0.951			
AU26			0.786			
Name:	Enjoyment Smile	Eyebrows Up	Mouth Open	Mouth Tightening	Eye Tightening	Mouth Frown

Figure 1. Factor enactments. A) Neutral face, B) F1: Enjoyment smile, F2: Eyebrows up, F3: Open mouth, F4: Mouth tightening, F5: Eye tightenting and F6: Mouth frown

were encouraged to express themselves and share feelings and personal experiences over a semi-structured interview for distress assessment. Finally, the *CRA* dataset consists of dyads engaging in a negotiation task over distributing a set of items. The large amount of data used as well as the different context and collection methods of the different datasets gives a context-free validity to those factors. All data described in Table 1 were analyzed with a commercial software based on CERT [12]. This software returns continuous evidence for individual AUs at 30FPS (a list of which can be seen in Figure 2). Facial AU evidence values were then resampled at dt=0.06s ($\sim$16 FPS) to smoothen out tracker artifacts and then thresholded to keep only active variance. The number of factors (n=6) was confirmed both by using the visual variance slope as well as the eigenvalue "greater-than-one" criterion.

The result of the factor analysis can be seen in Table 2, with a loading cutoff of 0.3. Some of these groupings make sense morphologically (E.g., Factor2 grouping together AU1 and AU2), however, in some other cases AUs from different part of the face are combined (like Factor1, combining AU6 and AU12). In that case, the new factor corresponds to a naturally occurring expression (E.g., Factor1 corresponds to a Duchenne Smile) that has psychological and social meaning. Based on literature we interpret those factors as follows:

F1   Enjoyment Smile - Associated with joy [25], [26]
F2   Eyebrows up - Conversational signal: used as emphasis mark / question mark [25]
F3   Open Mouth - Not typically associated with particular emotions on its own, unless combined with AU5 as surprise
F4   Mouth Tightening - Appeasement [10], [26]
F5   Eye Tightening - Anger [25], Goal obstruction [27]
F6   Mouth Frown - Unpleasantness [27]
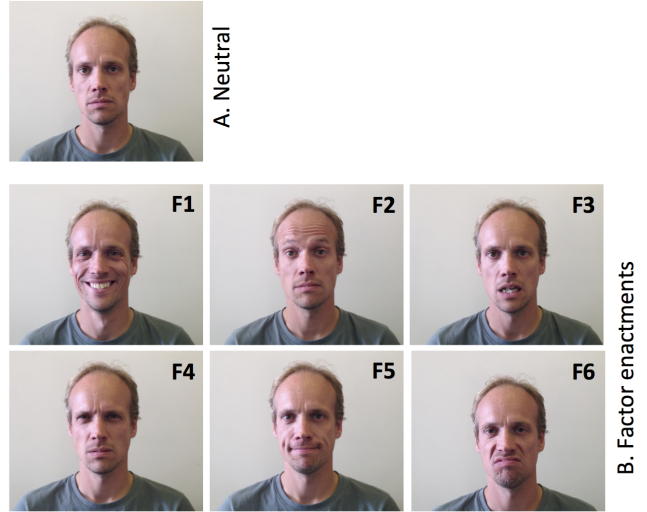
TABLE 3. Game Behavior Metrics from IPD

Metric Name	Description
pickC	Chance of player cooperating in any round
oppPickC	Chance of opponent cooperative in any round
#CC	Number of mutual cooperation states achieved in the game
#CD	Number of states where player was defected upon
#DC	Number of states where player defected on opponent
#DD	Number of mutual defection states in the game
#betrayed	Number of betrayals received by the player
#betray	Number of player betrayals of the opponent
#player_score	Overall player score at the end of the game
#opp_score	Overall opponent score at the end of the game
#dyad_score	Joint score at the end of the game
#score_diff	Score divergence at the end of the game

An enactment of the emerging factors can be seen in Figure 1.

In addition to the factor analysis conducted on the pooled datasets, we also conducted this analysis on each of the sets separately. The pattern of results was very consistent: in all three sets the same factors emerged, with only some minor variation between them. For example, the order of the factors was slightly different for the CRA dataset, but importantly the same factors emerged. Also, in the IPD and the CRA datasets AU20 was not present in the Open Mouth factor, whereas in both these sets AU10 contributed to the Eye Tightening factor. Details of the factors and their loadings when analyzing the datasets separately can be found in supplemental materials. The consistency in the results across
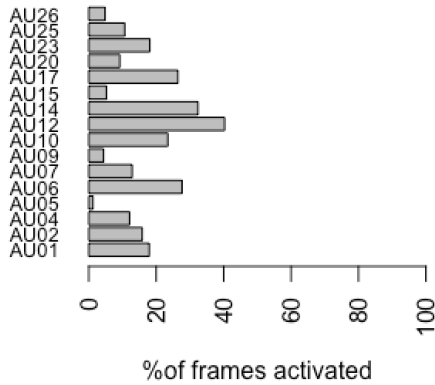
Figure 2. Facial action unit activation in the target analysis data, assessed automatically

these very different datasets speaks to the stability of the recovered factors and points to the generalisability of these factors in other social dilemmas and natural occurring social interactions.

4. Facial Expression Analysis in an Iterated Prisoner's Dilemma using Factors

4.1. Iterated Prisoner's Dilemma Corpus

For the Iterated Prisoner's Dilemma (IPD) task we obtained the data mentioned in relevant work [22]. This corpus includes videos of 608 participants (304 dyads) on an IPD task over 10 rounds, with game event annotations as to whether they cooperated with- (C) or defected upon- (D) their opponent.

Game Behaviors: Based on the IPD data we extract the following action metrics, or game behaviors as described in Table 3. [1] Those metrics capture elements of cooperative and non cooperative behavior, as well as joint game outcomes.

Facial Expressions: This data is challenging in terms of facial expression analysis. Participants were not allowed to talk to each other and were focusing on the game interface on the screen. On one hand, this enables facial expression analysis to be decoupled from speech facial movement; on the other hand, this configuration resulted in limited facial expressivity overall. As an overview, Figure 2 shows the percentage of frames where each AU is active (calculated by thresholding the continuous signal for positive evidence activation at 0.1). One may observe that only a few AUs occur beyond 20% of the frames in this data.

Correlation between the factors and the AU signals for the target analysis data is shown in Figure 3. This serves as a qualitative confirmation that the factors that were derived from general data still maintain meaningful association with AUs in the target dataset. We then proceed to test the external validity of those factors.

1. As a clarification: *DC* is the state where a player picks defect and the opponent picks cooperation. In this context, we call betrayal the transition to this state following a joint cooperation state (CC → DC)
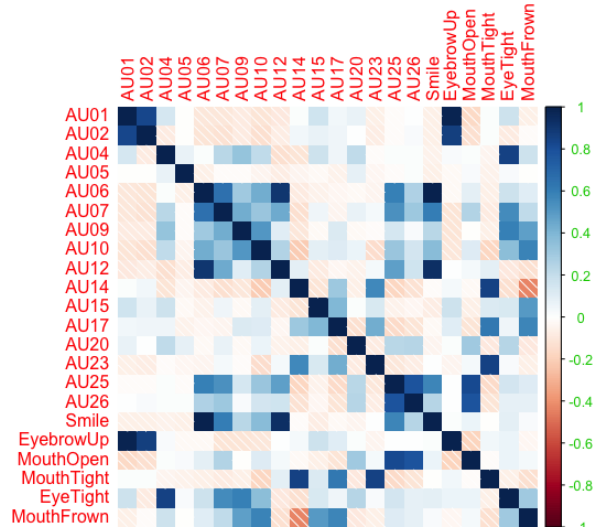

Figure 3. Correlation table between all signals in IPD data

4.2. Correlation with Game Behaviors

Here we look at facial expressions and how they associate with game behaviors in the IPD task. We look at both the traditional construct of facial AUs and the new factors we created, testing their external validity. We are looking at the whole interaction as a unit of analysis, and for this we summarize facial expression (by averaging) and game behavior measures over the 10 rounds.

Results are presented in Table 4, where one can see the significant correlation coefficients between game outcome (columns) and facial expression measures (rows).

There are some interesting observations to be made from this table: AU09 and AU10 are related to the same behaviors. The correlations suggest that nose wrinkles are associated with less prosocial choices (pickC), and also with less prosocial behavior of the opponent (less oppPickC, more mutual destruction). The outcome scores reveal that lower joint outcome increases nose wrinkling, or that higher joint outcome decreases nose wrinkling. Given the negative connotation that these expressions have [25], [28], this observation is in line with previous literature that negative expressions are associated with antisocial game acts and long term loss [22]. For AU14 (dimpler) the exact opposite pattern seems to emerge: Showing dimpler is associated with more pro-social choices (more pickC), and more pro-social behavior from the opponent (oppPickC). The outcome scores reveal that positive results for both the player and the opponent increases dimpler activation (or vice versa). It is noteworthy that AU14 (dimpler) is traditionally considered to have negative connotation (associated with contempt [29]) so its association with positive social outcome is interesting. For AU12 too, an interesting pattern of results emerges: AU12 is associated with pro-social choices (pickC) and mutual defection is associated with less smiling. However, getting betrayed is also associated with AU12 activity. This may mean that people smile away that they are being betrayed, but it could also reveal that although smiling associates with pro-social choices this makes people vulnerable

TABLE 4. OVERVIEW OF CORRELATIONS OF EXPRESSIONS WITH GAME BEHAVIORS, ****p < .0001, ***, p < .001, **,p < .01, * p < .05 , . p < .1

	pickC	oppPickC	#CC	#CD	#DC	#DD	#betrayed	#betray	pl_score	opp_score	dyad_score	score_diff
AU01				-0.07.	-0.07.							
AU02			0.07.	-0.11**	-0.07.		-0.07.					
AU04												
AU05	-0.07.	-0.07.				0.11**				-0.08.		
AU06							0.07.					
AU07												
AU09	-0.08*	-0.09*	-0.08*			0.1*				-0.07.	-0.11*	
AU10	-0.1*	-0.09*	-0.1*			0.09*				-0.08.	-0.11*	
AU12	0.08*	0.07.	0.07.			-0.08*		0.09*		0.08.	0.08.	
AU14	0.11**	0.13**	0.11**			-0.12**			0.11**	0.1*	0.16***	
AU15												
AU17				-0.09*	0.07.		-0.12**		0.09*	-0.1*		0.13**
AU20												
AU23		0.08.	0.08.	-0.08.								
AU25								0.09*				
AU26								0.13**				
EnjoySmile	0.07.					-0.07.	0.08.					
EyebrowUp			0.07.	-0.08*	-0.07.							
MouthOpen					0.07.			0.11**				
MouthTight	0.07.	0.12**	0.1*	-0.08*			-0.09*		0.11**		0.11**	
EyeTight												
MouthFrown	-0.16***	-0.11**	-0.13**		0.09*	0.13***	-0.13***			-0.17****	-0.15***	0.09*

for exploitation [15]. On a secondary note, there's a correlation between AU5 and DD, which may reveal that when people both defect they respond with fear [25] (perhaps fear for a downward spiral of defection). Finally, AU17 (chin raiser) is unexpectedly associated with betrayal (it correlates positively with DC state, increasing player score, increasing score divergence between players, decreasing opponent score, and less CD and getting betrayed states) and AU2 is unexpectedly negatively associated with number of CD.

Looking at the factors: F1-Smiling largely follows the same pattern as AU12 and so this can largely be interpreted in the same way: smiling is associated with pro-social choices, but this may make people vulnerable for exploitation [15] (or it could also mean that people smile when being exploited). F2 -Eyebrows up could be interpreted as a social marker that emphasizes particular behavior: In this case, when both parties cooperate people raise their eyebrows, perhaps encouragingly. When the opponent defects this is negatively related with this positive social marker (potentially discouraging this behavior). F3 -Mouth Open is associated with betraying the other player. Here we could speculate that people open their mouths to verbalize an apology, or as a submissive signal of fear [25] (expressing "oops"). F4 -Mouth tightening largely follows the same pattern as AU14. The unilateral occurrence of AU14 is sometimes associated with contempt [25], [28], but the factor suggests that in the context of multiple other AUs it is more in line with the interpretation that this signals appeasement / affiliation [10], [26]. F5 -Eye tightening is characteristically associated with anger and goal obstruction [25], [27], but is here somewhat surprisingly unrelated to any behavior. One possibility for this is that eye tightening is also associated with concentration [5], [30], [31], and so in the context where people are actively focusing on a computer screen, the expression may not have the angry connotation. F5 -Mouth Frown: Interestingly, the mouth frown seems

to combine the pattern of correlations that was observed for AU10 and AU17 separately. This helps explain the somewhat unexpected finding of AU17 correlations alone and places the finding for AU9/AU10 in relation to more than just isolated nose wrinkling/lip raising. As a factor these AUs are associated with less prosocial choices and also with less prosocial behavior from the opponent. The mouth frown expression is thought to communicate an unpleasant state and is associated with signals of appeasement such as sadness [25] and guilt/regret [16], but as a factor it is associated with less prosocial choices and more defection; so rather than restoring cooperation the consequence seems to be a breakdown of trust that leads to a decrease in opponent and dyadic scores.

It is particularly noteworthy that Mouth Tightening and the Mouth Frown have opposite effects. This suggests that friendly non-verbal communication (but not overtly smiling) can contribute to higher joint gains, whereas sulking negatively affects the game and decreases joint gains (possibly by contributing to a negative spiral). Interestingly, AU17 plays a role in both of those factors, but with different social outcomes, so the construction of factors demonstrates the benefit of investigating facial expressions in combination rather than in isolation.

4.3. Comparing Factor and AU contribution

Additional experiments on the prediction power of the factor set reveals that these groupings of AUs not only reduce complexity in a model, but also hold psychological meaning in that these are better predictors of choices and outcomes in the IPD task than the AUs alone. Table 5 shows the results of linear regression models that include only AUs (*out_AU*, with 16 degrees of freedom) and only Factors (*out_F*, with 6 degrees of freedom) as predictors of game behavior and outcome. The analysis reveals that both the AU model and the factor model contribute to the

TABLE 5. Contribution of AUs only and Factor only models in prediction of game behavior metrics

Behavior Metric	Model	Estimate	Std. Error	t value	Pr(>|t|)
pickC	out_AUs	-0.5613	0.3115	-1.80	0.0721 .
	out_F	1.2315	0.3902	3.16	0.0017 **
oppPickC	out_AUs	-0.1858	0.2748	-0.68	0.4991
	out_F	0.7655	0.3693	2.07	0.0386 *
#CC	out_AUs	-0.8634	0.3370	-2.562	0.010658 *
	out_F	1.4748	0.4199	3.512	0.000478 ***
#CD	out_AUs	-0.3987	0.3217	-1.24	0.2158
	out_F	0.8089	0.4383	1.85	0.0654 .
#DC	out_AUs	-0.8658	0.3434	-2.52	0.0120 *
	out_F	1.1994	0.4761	2.52	0.0120 *
#DD	out_AUs	0.1668	0.2294	0.73	0.4675
	out_F	0.4012	0.3331	1.20	0.2289
betrayed	out_AUs	0.0128	0.2734	0.05	0.9627
	out_F	0.6463	0.3670	1.76	0.0787 .
betray	out_AUs	0.1303	0.2416	0.54	0.5898
	out_F	0.0695	0.4083	0.17	0.8650
pl_score	out_AUs	0.0552	0.2543	0.22	0.8283
	out_F	0.3865	0.3844	1.01	0.3152
opp_score	out_AUs	-0.4410	0.3081	-1.43	0.1529
	out_F	1.1073	0.3940	2.81	0.0051 **
dyad_score	out_AUs	0.2075	0.2142	0.97	0.3332
	out_F	0.4290	0.3093	1.39	0.1660
score_diff	out_AUs	-1.2191	0.4063	-3.000	0.0028 **
	out_F	1.7566	0.5154	3.408	0.0007 ***

prediction of certain game behaviors (such as number of joint cooperation states, namely *#CC*), but that the factor model comes out as a better predictor for almost all metrics (in terms of t- and p-values). A Wilcoxon signed rank test comparing the absolute t-values of *out_AU* vs. *out_F* for all behavior metrics discussed in this paper shows that the factor model has significantly more effect than the AU model (Z = 3, p = .002). This indicates that the factors explain game behavior with more precision than their component parts.

5. Discussion

We presented an approach for facial expression analysis that falls between the extremes of low-level action units and high-level basic expressions. We performed a factor analysis of individual action units over several very large corpora of facial expressions, producing a set of 6 factors. The result yields interesting contributions for affective computing.

First, we showed that there is indeed reliable co-occurrence of certain AUs and we were able to extract six main factors. Interestingly, some of these correspond or have great overlap with facial expressions that have been suggested to be displays of basic emotion (like Enjoyment Smile and Eye Tightening). However, some factors that came up did not necessarily reflect basic emotions, although they did show resemblance to other known expressions (Mouth Tightening and Mouth Frown). These factors seem to capture facial expression variations with subtle distinctions (e.g., enjoyment smile, mouth tightening, mouth frowning) that occur in naturalistic interactions. The fact that

the same consistent pattern of factors emerges when we look at the different datasets individually speaks to the stability of these factors and hints to generalizability across different context.

Second, we show that these factors have external validity in that they are predictive of a person's behavior, the behavior of their opponent, and the outcome in an iterative social dilemma. The correlation with the IPD game behaviors can be interpreted in light of existing literature about the social meaning of these facial expressions [5], [10], [25], [26], [27]. Moreover, we claim that the extracted factors offer additional insights compared to the analysis of individual AUs.

Specifically, the factors seem to distinguish between different typologies of smile. Factor 1 captures joyful (Duchenne) smiles that usually associate with receiving or giving reward or enjoyment [26]. Interestingly, this smile is correlated with pro-social choices, but also with getting betrayed. As previously mentioned, this could mean that people "smile away" events [32], but could equally mean that smiling makes people vulnerable for exploitation [15]. The results reveal that it is not joyous smiles that mostly signal cooperation. Mouth tightening controls that are typically associated with affiliative smiles and are observed in social polite interactions [10], [26] are more predictive of pro-social outcomes than joyous smiles in the context of this social dilemma.

Our findings with regards to Eye Tightening may suggest that additional context need to be taken into account to reliably predict the effect of anger expressions in social dilemmas. Lastly, the factor analysis allows for a meaningful distinction between two morphologically close expressions: mouth tightening and mouth frown, that are associated with completely opposite social outcome in this scenario. The factor grouping even helps dissect the effects of certain AUs (like AU17) that shows up in both factors with opposite effects.

Although these observations are based on correlational rather than causal relations between facial and game behavior, they do support the notion that AUs dont occur in isolation. The morphological features of AUs change with the presence and absence of other AUs. Most importantly the presumed psychological meaning of AUs depends on their co-occurrence [25], which is captured in part with the factors. The results of this analysis are very encouraging that these factors have psychological meaning and can serve as units of analysis in facial expression studies. We believe that by re-shaping the constructs we use for analysis to better fit naturalistic data, we can overcome certain limitations in affective computing.

6. Conclusion

We performed factor analysis on a large data set of naturally occurring facial expressions and recovered psychologically meaningful factors that have been discussed in the literature. Further we showed that these factors have external validity in that they predict behaviors in a social

dilemma (IPD) and often with more precision than the AUs alone. Further work could validate the predictive value of the extracted factors in different context.

Acknowledgments

References

[1] P. Ekman, "An argument for basic emotions," *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.

[2] J. A. Russell, "Core affect and the psychological construction of emotion." *Psychological review*, vol. 110, no. 1, p. 145, 2003.

[3] A. J. Fridlund, "The new ethology of human facial expressions," *The psychology of facial expression*, vol. 103, 1997.

[4] P. Ekman and W. V. Friesen, "Facial action coding system," 1977.

[5] C. Darwin, "The expression of the emotions in man and animals john murray," 1872.

[6] P. Ekman, W. V. Friesen, and P. Ellsworth, *Emotion in the Human Face: Guide-lines for Research and an Integration of Findings: Guidelines for Research and an Integration of Findings*. Pergamon, 1972.

[7] H. Gunes and H. Hung, "Is automatic facial expression recognition of emotions coming to a dead end? the rise of the new kids on the block," *Image and Vision Computing*, vol. 55, pp. 6–8, 2016.

[8] N. Yitzhak, N. Giladi, T. Gurevich, D. Messinger, E. Prince, K. Martin, and H. Aviezer, "Gently does it: Humans outperform a software classifier in recognizing subtle, nonstereotypical facial expressions." *Emotion (Washington, DC)*, 2017.

[9] P. Robinson and T. Baltruaitis, "Empirical analysis of continuous affect," in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, Sept 2015, pp. 288–294.

[10] D. Keltner, "Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame." *Journal of personality and social psychology*, vol. 68, no. 3, p. 441, 1995.

[11] J. A. Russell and L. F. Barrett, "Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant." *Journal of personality and social psychology*, vol. 76, no. 5, p. 805, 1999.

[12] G. Littlewort, J. Whitehill, T. Wu, I. Fasel, M. Frank, J. Movellan, and M. Bartlett, "The computer expression recognition toolbox (cert)," in *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*. IEEE, 2011, pp. 298–305.

[13] T. Baltrušaitis, P. Robinson, and L.-P. Morency, "Openface: an open source facial behavior analysis toolkit," in *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*. IEEE, 2016, pp. 1–10.

[14] M. Bartlett, G. Littlewort, E. Vural, K. Lee, M. Cetin, A. Ercil, and J. Movellan, "Data mining spontaneous facial behavior with automatic expression coding," *Verbal and Nonverbal Features of Human-Human and Human-Machine Interaction*, pp. 1–20, 2008.

[15] G. A. Van Kleef, C. K. De Dreu, and A. S. Manstead, "The interpersonal effects of anger and happiness in negotiations." *Journal of personality and social psychology*, vol. 86, no. 1, p. 57, 2004.

[16] C. M. de Melo, P. J. Carnevale, S. J. Read, and J. Gratch, "Reading peoples minds from emotion expressions in interdependent decision making." *Journal of personality and social psychology*, vol. 106, no. 1, p. 73, 2014.

[17] D. Matsumoto, "Facial affect reciprocity in dyadic interactions," DTIC Document, Tech. Rep., 2012.

[18] G. A. Van Kleef, C. K. De Dreu, and A. S. Manstead, "An interpersonal approach to emotion in social decision making: The emotions as social information model," *Advances in experimental social psychology*, vol. 42, pp. 45–96, 2010.

[19] W. Poundstone, *Prisoner's Dilemma*, 1st ed. New York, NY, USA: Doubleday, 1993.

[20] C. Camerer, *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press, 2003.

[21] R. H. Frank, *Passions within reason: the strategic role of the emotions*. WW Norton & Co, 1988.

[22] G. Stratou, R. Hoegen, G. Lucas, and J. Gratch, "Emotional signaling in a social dilemma: An automatic analysis," in *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*. IEEE, 2015, pp. 180–186.

[23] J. Gratch, R. Artstein, G. M. Lucas, G. Stratou, S. Scherer, A. Nazarian, R. Wood, J. Boberg, D. DeVault, S. Marsella *et al.*, "The distress analysis interview corpus of human and computer interviews." in *LREC*, 2014, pp. 3123–3128.

[24] Z. Nazari, G. M. Lucas, and J. Gratch, "Opponent modeling for virtual human negotiators," in *International Conference on Intelligent Virtual Agents*. Springer, 2015, pp. 39–49.

[25] P. Ekman, W. Friesen, and J. Hager, "Facial action coding system: The manual on cd-rom. instructors guide," *Network Information Research Co, Salt Lake City*, 2002.

[26] M. Rychlowska, R. E. Jack, O. G. B. Garrod, P. G. Schyns, J. D. Martin, and P. M. Niedenthal, "Functional smiles: Tools for love, sympathy, and war." *Psychological Science.*, 2017.

[27] L. K. Pope and C. A. Smith, "On the distinct meanings of smiles and frowns," *Cognition & Emotion*, vol. 8, no. 1, pp. 65–72, 1994.

[28] P. Rozin, L. Lowery, and R. Ebert, "Varieties of disgust faces and the structure of disgust." *Journal of personality and social psychology*, vol. 66, no. 5, p. 870, 1994.

[29] P. Ekman and W. V. Friesen, "A new pan-cultural facial expression of emotion," *Motivation and emotion*, vol. 10, no. 2, pp. 159–168, 1986.

[30] P. Rozin and A. B. Cohen, "High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of americans." *Emotion*, vol. 3, no. 1, p. 68, 2003.

[31] P. C. Ellsworth, "Confusion, concentration, and other emotions of interest: Commentary on rozin and cohen (2003)." 2003.

[32] A. H. Fischer, D. Becker, and L. Veenstra, "Emotional mimicry in social context: the case of disgust and pride," *Frontiers in psychology*, vol. 3, 2012.