DOI: http://dx.doi.org/10.1145/2699255

The Effect of Wrinkles, Presentation Mode and Intensity on the Perception of Facial Actions and Full Face Expressions of Laughter

RADOSLAW NIEWIADOMSKI, DIBRIS, University of Genoa
and CATHERINE PELACHAUD, CNRS-LTCI, Telecom-ParisTech

This paper focuses on the identification and on the perception of facial action units displayed alone as well as the meaning decoding and the perception of full face synthesized expressions of laughter. We argue that the adequate representation of single action units is important in the decoding and the perception of the full face expressions. In particular we focus on three factors that may influence the identification and the perception of single actions and full face expressions: their presentation mode (static vs. dynamic), their intensity, and the presence of wrinkles.

For the purpose of this study we used a hybrid approach for animation synthesis that combines data-driven and procedural animations with synthesized wrinkles generated using a bump mapping method. Using such animation technique, we created animations of single action units and full face movements of two virtual characters. Next, we conducted two studies to evaluate the role of presentation mode, intensity and wrinkles in single actions and full face context-free expressions. Our evaluation results show that intensity and presentation mode influence 1) the identification of single action units, 2) the perceived quality of the animation. At the same time, wrinkles 3) are useful in the identification of a single action unit and 4) influence the perceived meaning attached to the animation of full face expressions. Thus, all factors are important for successful communication of expressions displayed by virtual characters.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Evaluation/ methodology*; H.1.2 [**Information Interfaces and Presentation**]: User/Machine Systems—*Human factors*; I.3.8 [**Computing Methodologies**]: Computer Graphics—*Applications*

General Terms: Human Factors

Additional Key Words and Phrases: facial expressions, action units, wrinkles, laughter

ACM Reference Format:

The Effect of Wrinkles, Presentation Mode and Intensity on the Perception of Facial Actions and Full Face Expressions of Laughter ACM Trans. Appl. Percept. 0, 0, Article 0 (2014), 20 pages.
DOI = 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

1. INTRODUCTION

The face is considered one of the most important channels of nonverbal communication. Facial expressions communicate a spectrum of emotions, underline attitudes, as well as serve regulatory functions

Author's address: R. Niewiadomski, DIBRIS, viale Causa 13, 16-145 Genoa, Italy; email: radoslaw.niewiadomski@dibris.unige.it; C. Pelachaud, 37/39 rue Dareau, 75014, Paris, France; email: catherine.pelachaud@telecom-paristech.fr

DOI 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

in conversations. Though there are still debates whether facial expressions encode internal states or serve purely communicative functions, the existence of facial expressions as social signals and their importance for communication is undoubted. Therefore, facial expression may also play an important role in a human-computer interaction, when it comes to the signalizing of emotions or states between the human and its virtual interaction partner, e.g., a *virtual character*. Virtual characters (VCs) are entities with a human-like appearance able to communicate verbally and nonverbally with human interlocutors. As a new paradigm of human computer interaction, VCs have found their way into several applications, e.g., the training of social and emotional skills (e.g. serious games), gaming industry, health and rehabilitation (e.g., personal assistants for elderly people), and other services (e.g., online shopping assistance, e-learning). Studying the perception of animated expressions is important, as correct communication of an emotion is a crucial aspect of VC applications. For example, the FearNot! application [Aylett et al. 2009] uses VCs to act out bullying situations with the aim to teach children coping strategies to handle bullying. In this application, the nonverbal communication of emotions by a VC, acting as the victim, makes children aware of the emotional states of the victims of bullying. But, human facial expressions are very complex activities that may involve large number of muscle contractions, consist of complex dynamics, and may involve physiological changes. Due to the complexity and variability of facial expressions and their subtle meanings it is still challenging to display a large spectrum of communicative intentions with synthesized expressions.

In this paper we study factors that influence the perception of facial expressions of a virtual character. Differing from most previous research, we focus not only on full face expressions but also on the way the local facial activity (called action unit (AU) [Ekman and Friesen 1975]) is displayed. Consistently with the recent research on emotion expressions in humans (e.g. [Scherer 2001], [Keltner 1995], see Section 2.1) we argue that also in the case of synthesized facial expressions, single action units may have an important role in the perception of the communicative intention transmitted with an expression. The best known (and most debated) example is the orbitalis oculi pars orbitalis activity (i.e., action unit 6 (AU6) in FACS; see [Ekman et al. 2002]) in the differentiation between the expression of amusement smiles and politeness or fake/posed smiles in humans [Ekman and Friesen 1975] and virtual characters [Ochs et al. 2012]. Sometimes, the misidentification of single action may even result in the misinterpretation of the whole expression. For example, confusing a polite smile with an amused smile might lead to the interpretation of a change of attitudes (e.g., positive feedback).

Most of the research on the identification and the perception of synthesized displays is restricted to six expressions of so called basic emotions [Ekman and Friesen 1975]. These expressions, important from an evolutionary perspective, are not the only ones to appear frequently in real life [Scherer et al. 2004]. Surprisingly, perception of other frequently used and significant expressions received much less attention from researchers. Laughter display is one of such expression. Laughter appears often in real life [Martin and Kuiper 2009] and it may have different meanings [Ruch et al. 2013]. Laughter facial displays are composed of several action units: some of them are obligatory; others are optional. Although the relation between different morphological configurations and perceived meanings remains in laughter yet unconfirmed - for the purpose of this study it is important that adding, removing or modifying the appearance of certain action units (i.e. components of the full face expressions) will result in displays that still can be recognized as a "plausible expressions of laughter". All these reasons make laughter expressions particularly suitable for the scope of this work.

In this paper we focus on three factors that may influence the perception of animated facial expressions. The first two factors: 1) the intensity of stimuli and 2) the presentation mode (i.e. whether a stimulus is presented statically as an image or dynamically as an animation) are often considered in studies on perception of human expressions (see Section 2.2 for details). We add a third factor: 3) the presence or absence of wrinkles. According to Hess and colleagues [Hess et al. 2012] wrinkles "can reinforce or obscure certain expressions by adding information to the face". For facial wrinkling, two types may be distinguished: age-related wrinkles and expressive wrinkles with two different roles in facial communication [Courgeon et al. 2009b]. The first type of wrinkles denominates a form of permanent age-related structural facial changes that is always visible. The second type appears and accompanies the activity of certain facial action units, typically linked to emotion or communicative signals. In this work we will focus only on the latter case. The role of expressive wrinkles in humans emotion communication was rarely considered in research, as such wrinkles are difficult to control in natural appearance. However, this difficulty of control can be overcome by using VCs. On one side VC technology permits us to fully control the creation of experimental stimuli and, thus, to prepare the pairs of animated expressions that can be differentiated only by presence or absence of wrinkles. On the other side, aiming to build a VC able to communicate emotions we need to understand how the expressions of VCs (with or without wrinkles) are perceived by the human interaction partners. In this paper, we are particularly interested in checking whether adding expressive wrinkles is only an arbitrary, "artistic" choice of a person who builds a virtual model or whether it has more significant consequences e.g., it influences the perceived meaning of the expression.

This work is composed of two perceptive studies showing the role of expression intensity, presentation mode, and wrinkles of animated facial expressions on their identification, perceived meaning, and animation quality. First, we measure the identification rate and animation quality of single action units (e.g., frowning, which is equivalent to AU4 in the FACS or lip corner raising, AU12) in a context free experiment. Still, action units are often elements of full face expressions. Thus, the second study focuses on full face expressions of laughter with variable composition of action units, and we measure their perceived meaning and animation quality. The next section contains a survey of works on the identification of human and VC facial expressions. Section 3 is dedicated to the animation synthesis approach we use in this work. Section 4 presents the study on identification and perception of single action units while Section 5 presents the results of the evaluation of full face expressions of laughter. We present a general discussion in Section 6 and we conclude the paper in Section 7.

2. BACKGROUND

2.1 Role of Single Action Units in Human Facial Expression Decoding

Several findings exemplify the role of local facial activity, i.e. action units for the identification and perception of a full face facial expression. Many researchers (e.g. [Scherer 2001; Kaiser and Wehrle 2001]) share the opinion that facial expressions are visible evidence of appraisal processes. Consequently, single components like raising the eyebrows or opening the mouth may have an independent meaning. These facial actions indicate the specific appraisal results e.g. (high) suddenness corresponds to raising the brows (AU1 + AU2) and parting the lips (AU25) [Kaiser and Wehrle 2001]. Within research on human facial expressions, probably the most well examined example of a local facial cue, which can alter perception of whole expression, is the orbitalis oculi pars orbitalis activity (i.e., action unit 6, AU6). AU6 raises the cheeks and compresses the eyelids due to the muscle contraction. This movement pulls the skin toward and around the eyes, creating "crows-feet" wrinkles and is typically involved in the prototypical expression of joy, the felt smile [Frank and Ekman 1993], but does not necessarily occur in other forms of smiles, like polite or fake smile [Ekman and Friesen 1975]. The role of other action units in distinction of different communicative intentions expressed with smile (i.e., embarrassment and amusement) was showed by Keltner [Keltner 1995].

2.2 Role of Wrinkles, Presentation Mode and Intensity in the Identification of Humans' Expressions

Within the analysis of human facial expressions, Ekman and colleagues have developed the Facial Action Coding System [Ekman et al. 2002]. For each action unit, the authors have defined guidelines how to detect it on five different levels of intensity, coded from trace (coded as A) to maximum (E, no further stretching, bulging, pouching is possible). In their descriptions, wrinkles play an important role in the adequate identification of AUs and their intensity. Consequently, FACS coders rely on wrinkles: their appearance, intensity and shape when coding facial expressions. For instance the specific wrinkles are an indicator of middle and intense AU6 (intensity B or more); the intensity of these wrinkles is used to distinguish between AU6 and AU7 (see FACS manual [Ekman et al. 2002]).

Hess et al. [Hess et al. 2012] studied the role of age-related structural changes on the face such as wrinkles and folds in the nonverbal communication of sadness and happiness. According to their results, age-related wrinkles reduce the signal clarity; emotions were perceived less accurately when shown by an old face. Additionally, the correct (i.e., intended by the displayer) emotions were perceived as less intense, while incorrect ones as more intense in old faces. Thus, the permanent age-related wrinkles produce the expressions that are perceived as displays of "mixed emotions". They also found that faces of elderly persons showing neutral expression are perceived as expressing some emotions at a higher rate than faces of young people. Thus, age-related wrinkles can be a cue to perceive expression of emotion even when no emotion is being displayed on the face.

Multiple experiments show that the presentation mode (i.e. image vs. video) influences the recognition of emotional expression. Among others, the different reactions to dynamic and static facial expressions were observed by measuring EMG activity [Rymarczyk et al. 2011]. Bassili [Bassili 1979] showed that six basic emotions are better identified from dynamic stimuli. Similar effect was observed for the recognition of five basic expressions with low intensity [Ambadar et al. 2005]. Interestingly in the latter study, the happiness expression was an exception: it was correctly identified even from low intensity images. More recently, Krumhuber et al. [Krumhuber et al. 2009] have shown that also temporal qualities of an expression (i.e. onset, offset duration) displayed by humans and virtual characters may influence the perception of the meaning of an expression.

Regarding intensity factor, Hess et al. [Hess et al. 1997] have showed that, in the case of four basic emotions (sadness, happiness, disgust, anger), the perceived intensity of expressions presented on images was correlated with their physical intensity. Additionally, in the case of three emotions (i.e., except happiness) emotion decoding accuracy varied linearly with the physical intensity. In the experiment of Young et al. [Young et al. 1997] morphed images of basic emotions and neutral expressions were used in an identification task. The images presented five degrees of the intensity from the neutral to the very intense expression. The results showed a high identification rate only for the images placed on the extreme points of the intensity continuum. A similar animation technique was used by Sato and Yoshikawa [Sato and Yoshikawa 2004] who showed that the perceived naturalness of the morphed videos (from neutral to full intense basic emotion) depends on the velocity of the presentation. Finally, Rymarczyk et al. [Rymarczyk et al. 2011] showed that human dynamic expressions were rated as more intense than static ones.

2.3 Laughter Expressions

Laughter may serve several conversational functions; it is also a prominent behavioral marker for emotions (i.e., amusement, joy). Although there is no agreement yet on the number of laughter types, the minimal distinction will be, as Ruch and Ekman [Ruch and Ekman 2001] suggested, the differentiation between spontaneous and fake laughter. According to the literature, all laughter expressions contain the activation of the zygomatic major muscle that corresponds to the action unit AU12. Additionally the cheek raise is often present (AU6). Its presence is associated with the Duchenne or joyful laughter [Ruch 1993]. Different authors [Keltner 1995; Ruch 1993] defined the basis of joyful/amused laughter consisting of the Duchenne Display (AU12 + AU6) and an open mouth (AU25, and maybe also a jaw relaxation/drop: AU26/27). Also the inner part of the orbicularis oculi muscle (AU7) was found to be frequently present in joy expressions. Darwin [Darwin 1872] suggested that intense laughter expressions include lowering of the eyebrows (AU4). In the acted expressions of hilarious laughter, the actions AU25, AU26 are frequently observed while AU7, AU27 (Mouth Stretch), AU4 (Brow Lowerer) occur less and AU1 (Inner Brow Raise), AU2 (Outer Brow Raise), AU9 (Nose Wrinkler) and AU20 (Lip Stretcher) occur occasionally [Drack et al. 2009]. In naturally occurring laughter, actions AU5 (Upper Lid Raiser), AU6, AU7, and AU9 are observed [Ruch and Ekman 2001].

2.4 Role of Single Action Units in Virtual Characters Expressions

The role of single actions on the global meaning of synthesized facial expressions was recently explored. Inspired by the Scherer's appraisal theory [Scherer 2001], Courgeon and colleagues [Courgeon et al. 2009b] built a model in which the generation of facial expressions is directly driven by the evaluation of events appraised by a VC. The expressions in this model are generated at two levels. First, a temporary animation, corresponding to the currently evaluated appraisal variable is displayed. Second, when the evaluation of the event through all appraisal variables is finished, the system computes to which emotion the sequence of appraisal values corresponds and displays the corresponding full face expression. Complementary work on the interpretation of the facial behavior of a virtual character was proposed by de Melo et al. [Melo et al. 2012]. According to this approach, called “reverse appraisal”, the social impact of a character's expression is not related to the expression itself but to the observer's perception of how the VC was appraising the event, as well as its intentions. de Melo et al. [Melo et al. 2012] showed that people in a negotiation scenario were able to retrieve information from the character's facial expressions about how the goal conduciveness and blamefulness were appraised by the virtual character. While not directly related to laughter, an interesting work on the role of smiles were proposed by Ochs et al. [Ochs et al. 2012]. They proposed an algorithm to generate different smiles for three different smile meanings: a smile as sign of amusement, embarrassment, or politeness. In more details, they used machine learning technique to build a decision tree, which is used to choose between several morphological (i.e., several action units) and dynamic features (i.e. duration of the expression, onset and offset) of the smile expression.

2.5 Role of wrinkles, presentation mode and intensity in the identification of virtual characters expressions

Katsyri and Sams [Katsyri and Sams 2008] showed that synthesized dynamic expressions of basic emotions are identified better than static ones for expressions whose static displays were not-distinctive. In a similar study of Noël et al. [Noël et al. 2006] the effect of the presentation mode was not observed. In Bartneck and Reichenbach's work [Bartneck and Reichenbach 2005] the higher intensity expressions were better recognized compared to low intensity ones. In Courgeon et al. [Courgeon et al. 2009a] the application of wrinkles increases the character's expressivity but does not improve the recognition of basic emotions. Compared to static images, dynamic morphed expressions of three displays (neutral, happy, angry) led to better recognition rates and higher intensity and realism ratings in the study of Weyers et al. [Weyers et al. 2006].

3. LAUGHTER STIMULI GENERATION

For the purpose of our experiment we enhanced a freely available VC with the capacity of laughing using a hybrid approach that mixes two different animation techniques i.e., procedural animation and data-driven approach. The reasons that we merged these two techniques are the following:

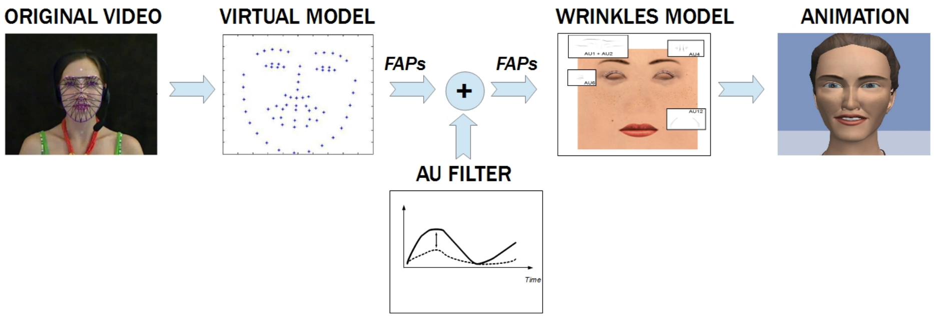

Fig. 1. Laughter animation synthesis.

(1) Data-driven based animation is richer in movements and consequently it may be perceived as more realistic. It permits maintaining the temporal and dynamic characteristics of the original expression. However, animation generated with this method is difficult to control and manipulate precisely (e.g., its duration, intensity, appearance of AUs);
(2) Specifying procedurally sequences of action units allows one to control facial animation and its meaning precisely (e.g., by adding or removing some actions such as AU6) but has all the weaknesses of procedural animation. The limitations entail that the animation is poor in details and the dynamics of the movements are not very realistic.

For the purpose of this study we used a virtual character called Greta [Niewiadomski et al. 2011] that can be driven by facial behavior description based on FACS [Ekman et al. 2002] or facial animation parameterization (FAPs) that is part of the MPEG-4 standard [Ostermann 2002].

3.1 Dataset

To generate the stimulus material we used the freely available Audio Visual Laughter Cycle (AVLC) corpus [Urbain et al. 2010] that contains spontaneous audio-visual laughter episodes with no overlapping speech. Each subject was recorded watching a 10-minutes comedy video. Participants' facial expressions were captured with a motion capture system and synchronized with the corresponding audiovisual sample. The material was manually segmented into episodes containing just one laugh. Next, through perceptive study, coders annotated the perceived intensity of the AVLC episodes using a Likert scale from 1 (low intensity) to 5 (high intensity) (for details see [Niewiadomski et al. 2012b]).

In this work we used the data of one female participant (subject number 5 in the corpus).

3.2 Hybrid Animation Synthesis

The synthesis method used to generate the stimuli of the experiments consists of three steps (see Figure 1): 1) retargeting the motion capture data, 2) applying AUFilters that, by applying a procedural method, modulate the animation of certain action units, 3) adding the wrinkles. In more details, the laugher animations generated for the purpose of this study combined different animation techniques developed previously [Niewiadomski et al. 2012a; Qu et al. 2012]. The animation generation was mainly based on the data-driven approach proposed by Qu and colleagues [Qu et al. 2012]. In this approach, the videos of the human are processed with a facial tracker; the tracked data are consequently mapped onto the mesh of a virtual character. In the second step, the result of data-driven approach was modified by applying AUFilters to some parts of the face. Both input and output data of AUFilters are expressed in FAPs but an AUfilter acts at the level of AUs and not of FAPs level as AUs are the smallest biological units. FAPs were designed for the sole purpose of building animation system. Thus, AUfilter modifies certain FAPs that corresponds to one AU. For instance, in order to add the activity of AU6 to an existing animation, the values of FAPs simulating the AU6 activity on Greta's face are generated with the procedural approach while the rest of the animation is generated by the data-driven approach. Using this hybrid method we have a better control over the expression; for instance, we can add or remove some actions units but we maintain the global dynamics of the original expression. In the third step of animation generation, the wrinkles were added. Each wrinkle in our model is associated with the activity of a specific action unit. The appearance and intensity of a wrinkle depends on the AU intensity. More details are presented below.

3.2.1 *Data-driven synthesis.* In order to estimate FAPs of natural facial expressions, Qu and colleagues [Qu et al. 2012] used an open-source face tracking tool - FaceTracker [Saragih et al. 2011] - to track facial landmark localizations. It uses a Constrained Local Model (CLM) fitting approach that includes a Regularized Landmark Mean-Shift (RLMS) optimization strategy. It can detect 66 facial landmark coordinates within real-time latency depending on the system's configuration.

To compute FAPs from facial landmarks, the authors created a neutral face model from 50 images of different persons' neutral faces. Next, FAPs are estimated from the distance between current facial landmarks and neutral face landmarks. Such mapping has to be created manually for each subject and VC (see [Qu et al. 2012] for details).

3.2.2 *Procedural synthesis.* To enhance data-driven animation with manually specified action units, we mapped each AU to FAPs, and then we provided, for each concerned AU, the procedures that modify its FAPs values. To avoid the abrupt changes or inconsistent animation on different parts of the face the change of the FAP values of modified AUs is proportional to the intensity of the activation of other facial actions presented in the original expression. Thus the activation of FAPs that correspond to modified action units is proportional to the activation of the FAPs of the original expression. For example, to add the AU4 to an existing animation, the values of the vertical displacement of left inner corner of eyebrows (FAP 4.2.x, see [Ostermann 2002]) are computed as:

$$\begin{array}{l} FAP\_4.2.x_{new} = FAP4.2.x_{old} * (1 + a) \\ if\ (FAP\_8.3.y > th_1)\ then\ a = 1\ else\ a = \frac{FAP\_8.3.y}{th_1} \end{array} \tag{1}$$

where $FAP\_8.3.y$ is the value of the horizontal displacement of the left mouth corner and th_1 is an empirically chosen threshold. Similar formulas are applied to the right side of the face. Then, adding the AU6 to an existing animation the values of the horizontal displacement of left cheek corner (FAP 5.3.y corresponding to AU6) are computed:

$$if\ (FAP\_8.3.y > th_2)\ then\ FAP\_5.3.y_{new} = (FAP\_5.3.y_{old} + FAP\_8.3.y) * th_3 \tag{2}$$

where $FAP\_8.3.y$ is the value horizontal displacement of left mouth corner and th_2 and th_3 are two thresholds.

This approach ensures that the modified facial actions will not appear abruptly and that the dynamics of the whole expression is maintained. The animation starts from the neutral expression, and enhanced facial actions are added gradually; they display the highest intensity in synchrony with the remaining parts of the face.

3.2.3 *Wrinkles.* As the main focus of this work is not related to the evaluation of any particular highly realistic wrinkles synthesis algorithm we used a standard approach characterized by low computation costs, i.e. the texture based technique called the "bump mapping" that was firstly introduced by Blinn [Blinn 1978]. This method consists in modifying the surface normal vector before the lighting computation, thus the bent normal can give the visually satisfying result without changing the surface

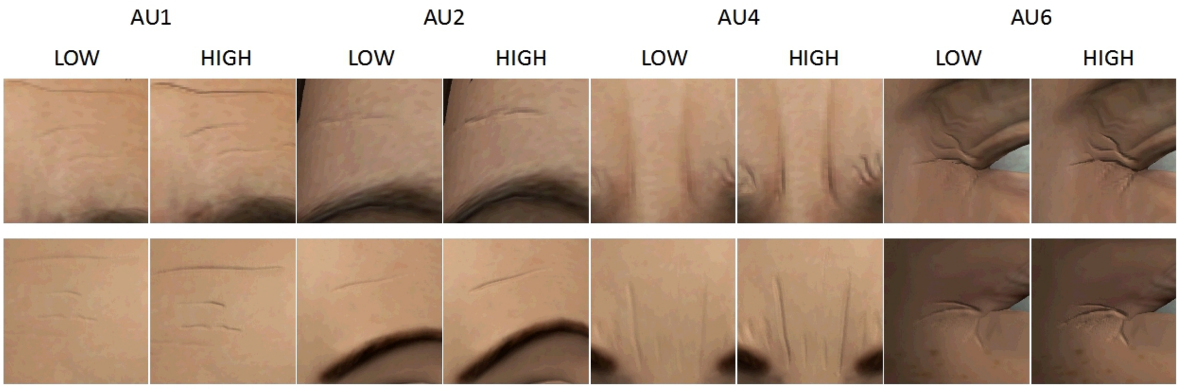

Fig. 2. Wrinkles associated with AU1, AU2, AU4, and AU6. First column - male character, second column - female character.

geometry. Groups of wrinkles, which correspond to different AU activation, were defined in textures of each virtual character. In runtime, when an action unit is to be activated on the face mesh, the graphic engine receives the AU intensity value and computes the intensity of corresponding wrinkles. The final animation (see Figure 2) is the composition of all active wrinkles. Details of this approach can be found in [Niewiadomski et al. 2012a].

4. STUDY 1: IDENTIFICATION, NATURALNESS AND REALISM OF SINGLE ACTION UNITS

In this first evaluation we studied factors that may influence the identification of single facial actions (e.g., frown or cheek raising). Previous studies on the perception of VC expressions used only the full face expressions of the basic emotions (see Section 2.5). To our knowledge there is no study that focuses on the identification of single action units. It is rather surprising because the appearance of even one additional action unit may change decoding of a full face expression (see Section 2.1). Single action units can be the components of different facial expressions. For example, the action: "lip corners pulled up" (i.e., AU12) occurs in the expression of joy as well as of embarrassment [Keltner 1995]. AUs also are more subtle (consequently they may be more difficult to identify). Taking into account the existing literature on identification of basic expressions, in the first study we tested three factors that may influence their perception: presentation mode (dynamic vs. static displays), intensity, and the role of wrinkles. We measured their identification rate, their realism and their naturalness.

4.1 Experiment setup

4.1.1 *Stimuli selection.* In the first study we evaluated the identification of single action units of laughter. Taking into consideration the list of action units observed in laughter we chose the following three actions: frown (AU4), cheeks raise (AU6), and lip corners pulled up (AU12). These actions are particularly important in laughter displays (see Section 2.3). They are very different from each other so we assume there would not be much confusion among them in the identification task. To increase ambiguity in the identification task we intentionally added three frequently used facial actions: oblique eyebrows (AU1+4), raised eyebrows in a half circle (AU1+2) and lip horizontally extended (AU20). We divided these six actions into two sets: upper-face (frowning, oblique eyebrows, raised eyebrows in half circle) and lower-face (cheeks raise, lip corners pulled up, lip horizontally extended, though we know that the AU12 may also lead to changes in the upper face).

The experiment entailed 96 different stimuli. Two different VCs were used: one female and one male. All six facial actions were displayed with (WR condition) and without wrinkles (NW condition) and with low and high intensities (see Figure 2). Each of the used action units (AU1, AU2, AU4, AU12, AU20) triggers a distinct wrinkle. Finally two different presentation modes were used: the participants were shown an image of the expression at its apex (static condition) or an animation of the facial action appearing and disappearing (dynamic condition), leading to a total of 96 stimuli. A certificated FACS expert validated all of the animations.

4.1.2 *Procedure.* The evaluation study was online as a set of web pages. Each web page displayed one stimulus. Participants could see the animations only once and they had to answer all questions before being allowed to see the next animation. The animations duration lasted 5 seconds while image displaying duration was not restricted.

Each participant could evaluate a balanced set of maximum 32 stimuli. It consisted of 8 upper and 8 lower-face animations as well as 8 upper and 8 lower-face images that were chosen randomly from a initial set of 96 stimuli. Participants were told they could stop the experiment wherever they wanted. This explains the different number of answers for each stimulus.

4.1.3 *Evaluation questionnaire and hypotheses.* During the experiment participants were asked to identify stimuli using 5-point Likert scale from “totally disagree” to “totally agree”. For the upper-face actions there were three separate scales labeled “raised eyebrow in half circle”, “frowning” and “raised eyebrow in oblique”. For the lower-face expressions there were three other scales: “cheek raise”, “lip extension - corner up”, “lip extension - straight”.

Intentionally we did not want to use neither the official action units names nor the muscle names of the corresponding actions replacing them with colloquial names understandable for non-expert (naive) participants. Moreover, in this study we avoided asking explicitly if the people could see any wrinkles. Thus some of the widely used terms were not appropriate e.g., “crows-feets” because these are explicitly used to describe the wrinkles and not a muscle action. For this reason we had to use less precise terms e.g., “cheek raising”. Participants were also asked to express their opinion on the naturalness and realism using two separate 5-point Likert scales. Five hypotheses were tested:

—H1.1. High intensity actions are better identified than low intensity ones;

—H1.2.a. Actions displayed by an animation (i.e. presentation mode: dynamic) are better identified than the ones presented by images (i.e. presentation mode: static);

—H1.2.b. Actions displayed by a dynamic animation are more natural and more realistic than the ones presented by static images;

—H1.3.a. Actions with wrinkles are better identified than actions without wrinkles;

—H1.3.b. Actions with wrinkles are more natural and more realistic than actions without wrinkles.

Regarding the hypotheses H1.1 and H1.2.a we expected that the factors: presentation mode and intensity will play similar role to the one observed for full face expressions of basic emotions in humans and virtual characters (see Section 2.2 and 2.5). Additionally animations of single facial actions (dynamic presentation mode; hypothesis H1.2.b) will be considered more natural (comparing to their opposites) as they naturally occur in the real-life. Regarding the hypotheses H1.3.a and H1.3.b it was assumed that synthesized facial wrinkles will play an important role in the decoding of synthesized AUs as it is in the case of humans’ displays. They also improve the perception of naturalness and realism of the expressions.

4.2 Results

In total, a sample of 1570 responses from 95 participants from 26 countries (63 males, age 19-51; mean age 27.4, SD = 4.4, most frequent countries of origin: France 28% Poland 11%, India 8% and Holland

Table I. Global effect on Naturalness and Realism.

	intensity				presentation				character				wrinkles			
	high		low		animation		image		male		female		no		yes	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
NAT	3.05	1.25	3.41	1.12	3.38	1.14	3.04	1.25	3.16	1.22	3.30	1.18	3.20	1.21	3.26	1.97
REA	3.03	1.25	3.39	1.10	3.35	1.23	3.04	1.24	3.14	1.20	3.29	1.17	3.21	1.20	3.22	1.18

6%) was collected. Due to the long duration of the experimental session many of the participants did not evaluate all 32 stimuli randomly chosen for them. Thus we decided to consider also all partial answers.

4.2.1 *Naturalness and realism.* First, Pearson correlations between Naturalness and Realism computed for the action types separately revealed high positive correlation ($r = .88$, $p < .001$).

Next, a MANOVA was conducted with the character Version (female, male), Presentation (dynamic vs static), Intensity (low vs. high) and Wrinkles (yes vs. no) as independent variables and Naturalness and Realism as dependent variables. The results (see Table I) show that the main effect of Version ($F(2, 1431) = 3.23$, $p < .05$; Wilk's $\lambda = .996$), Presentation ($F(2, 1431) = 15.27$, $p < .001$; Wilk's $\lambda = .976$), and Intensity ($F(2, 1431) = 18.79$, $p < .001$; Wilk's $\lambda = .974$) were significant. No interaction effects were observed.

In more details, the dynamic presentations were perceived as more natural ($F(1, 1432) = 30.51$, $p < .001$) and realistic ($F(1, 1432) = 25.05$, $p < .001$) compared to the static images. The low intensity animations were perceived as more natural ($F(1, 1432) = 35.48$, $p < .001$) and realistic ($F(1, 1432) = 35.47$, $p < .001$) than high intensity stimuli. The difference between two characters was also observed, the "female" character receiving higher scores for naturalness than the "male" one ($F(1, 1432) = 5.89$, $p < .05$). At the same time no significant effect of the character version on perceived realism was observed (applying Bonferroni correction, $F(1, 1432) = 6.26$, $p = .015$). Finally, there was no difference in perception of naturalness ($F(1, 1432) = .586$, $p = .44$) and realism ($F(1, 1432) = .01$, $p = .99$) regarding the use of wrinkles.

We additionally checked if a relation between the intensity of an action and its perceived naturalness and realism does not depend on the quality of wrinkles animation. The lack of interaction between Intensity and Wrinkles suggests that this effect does not depend on presence of wrinkles. To confirm this we ran a MANOVA only on animations without wrinkles; the results show similar effects of Intensity on Naturalness ($F(1, 720) = 30.996$, $p < .001$) and Realism ($F(1, 720) = 30.58$, $p < .001$).

4.2.2 *Identification task.* To investigate the effects of wrinkle presence, presentation mode and action intensity on the perception of the occurrence of each respective action, we considered each facial action separately. To test our three postulated hypotheses (H1.1, H2.1.a, H1.3.a), we used 3-way (2×2×2) ANOVAs with Wrinkles (yes vs. no), Presentation (dynamic vs. static), and Intensity (low vs. high) as independent variables (see Table II and Figure 3 for detailed results).

AU4. The results of the 3-way (2×2×2) ANOVA showed a significant main effect of Intensity ($F(1, 247) = 10.803$, $p = .001$) as well as of Wrinkles ($F(1, 247) = 62.299$, $p < .001$), and of Presentation ($F(1, 247) = 14.840$, $p = .001$) on identification rate (statement: "frowning").

Wrinkles × Presentation × Intensity interaction ($F(1, 247) = .760$, $p = .384$) was not observed. Still, the 2-way Wrinkles × Presentation interaction ($F(1, 247) = 19.412$, $p < .001$) was significant but the other interactions: Wrinkles × Intensity ($F(1, 247) = 2.798$, $p = .096$) and Intensity × Presentation ($F(1, 247) = .9114$, $p = .341$) were not significant. Next, the Wrinkles × Presentation interaction effect was analyzed using a post-hoc LSD test. The Presentation was significant in the no wrinkles (NW) condition ($F(1, 247) = 34.205$, $p < .001$) but not in the wrinkles (WR) condition ($F(1, 247) = .153$, $p = .696$).

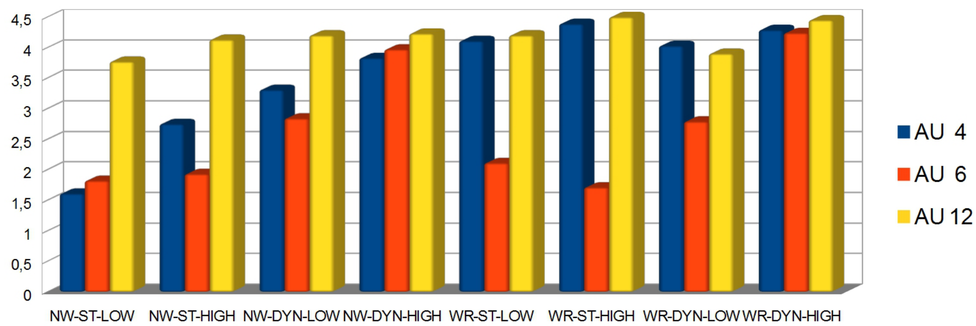

Fig. 3. Mean identification rates of three action units. Notation: WR - wrinkles stimulus, NW - without wrinkles stimulus, ST - static presentation mode (image) stimulus, DYN - dynamic presentation mode (animation) stimulus, LOW - low intensity stimulus, HIGH - high intensity stimulus.

The factor Wrinkles (present or not) influenced the identification of expressions from images (static condition; $F(1, 247) = 67.438$, $p < .001$) and animations (dynamic condition; $F(1, 247) = 6.921$, $p < .05$).
AU12. The results of the 3-way (2×2×2) ANOVA showed a significant main effect of Intensity ($F(1, 274) = 5.124$, $p < .05$) but not of Wrinkles ($F(1, 274) = 1.752$, $p = .187$) nor of Presentation ($F(1, 274) = .104$, $p = .747$). High intensity AU12 stimuli were better identified than low intensity ones.
Wrinkles × Presentation × Intensity interaction ($F(1, 274) = 1.168$, $p = .281$) was not observed. Also Wrinkles × Intensity interaction ($F(1, 274) = .703$, $p = .403$), Wrinkles × Presentation interaction ($F(1, 274) = 2.521$, $p = .113$) and Intensity × Presentation interaction ($F(1, 274) = .26$, $p = .871$) were all not significant.
AU6. The results of the 3-way (2×2×2) ANOVA showed a significant main effect of Intensity ($F(1, 257) = 10.584$, $p = .001$) and of Presentation ($F(1, 257) = 79.652$, $p < .001$) but not of Wrinkles ($F(1, 257) = .167$, $p = .683$).
Wrinkles × Presentation × Intensity interaction ($F(1, 257) = 1.328$, $p = .25$) was not observed. Intensity × Presentation interaction ($F(1, 257) = 16.532$, $p < .001$), was significant, but Wrinkles × Intensity ($F(1, 257) = 0.73$, $p = .788$) and Wrinkles × Presentation ($F(1, 257) = .42$, $p = .837$) interactions were not significant. Next, the Intensity × Presentation interaction effect was analyzed using a post-hoc LSD test. Intensity influenced identification from animations (dynamic condition; $F(1, 257) = 31.687$, $p < .001$) but not from images (static condition; $F(1, 257) = .286$, $p = .593$). However, Presentation mode influenced the identification of both low ($F(1, 257) = 11.836$, $p = .001$) and high ($F(1, 257) = 84.150$, $p < .001$) intensity stimuli.

4.3 Discussion

In this experiment we tested three action units: **AU4** (i.e. frown) is the only action out of three that in the first evaluation was better identified with wrinkles, from the animations (dynamic stimuli) and with higher intensity. The interaction effect between Wrinkles and Presentation shows the important

Table II. Identification of facial actions.

	intensity				presentation				wrinkles			
	high		low		animation		image		yes		no	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
AU4	3.84	1.46	3.31	1.64	3.86	1.50	3.21	1.59	4.20	1.26	2.96	1.40
AU6	3.14	1.66	2.46	1.62	3.46	1.69	1.89	1.15	2.83	1.67	2.76	1.69
AU12	4.33	1.08	4.02	1.20	4.20	1.18	4.15	1.11	4.25	1.09	4.10	1.21

role of expressive wrinkles. For wrinkles-based animations, the presentation mode did not influence the identification rate anymore. Thus, when wrinkles were present, even images were sufficient to signal the facial action. Adding wrinkles is particularly useful in the case of static stimuli containing AU4 i.e., when the movement of the brows from the neutral position cannot be seen (and the position of the brows in AU4 cannot be compared to their position in a neutral expression). In that case, wrinkles are the most important source of information allowing the observer to infer the displacement of the brows. In the dynamic condition the movement itself can be observed giving the same information to the observer. Consequently, the effect of wrinkles is still observed, but it is lower than in a static condition. **AU6** was identified much better in the intense and dynamic animations. Indeed, this action is particularly difficult to be perceived from the static images. The interaction effect observed between presentation mode and intensity of the stimulus suggests that displaying the movement (i.e. dynamic condition) is particularly useful for identifying AU6 with high intensity while it is less helpful in identifying AU6 with low intensity (in that case the identification was particularly low, independently of the presentation mode). **AU12** received very high identification rate in all the conditions. Neither wrinkles nor presentation mode improved the identification of this action. This observation is consistent with the results of Hess et al. [Hess et al. 1997] and Ambadar et al. [Ambadar et al. 2005] on the perception of human expressions showing that happiness expression is easily recognized even from low intense stimuli and/or images. Indeed, AU12 is a crucial cue of the expression of happiness.

Consequently, three out of five hypotheses were confirmed. The hypothesis H1.1 concerning the intensity of AUs was confirmed. The significant effect of intensity on the identification rate was observed for all three action units. The more intense AUs are better identified than the low intense ones. Interestingly they are perceived less natural and less realistic. Consequently, when creating animations of single action units animators need to balance between their readability and naturalness. Regarding the hypothesis H1.2.a, as expected, the animations were better identified than the images. The significant effect of presentation mode was observed for 2 out of 3 action units. This result confirms that the dynamics of an expression is an important cue allowing for correct identification. Animations contain all information about the movement and therefore outperform images when representing facial actions. Consequently, in the second evaluation we decided to use only dynamic presentation.

The effect of wrinkles on identification rate was limited (hypothesis H1.3.a). Only for one action the identification was significantly better compared to the no wrinkles condition. This result might be partially influenced by the use of imprecise labels to describe the action. In particular in the case of AU6 - the label used in the evaluation was “cheek rising” and not “crows-feets”. The latter does explicitly refer to wrinkles and not to the action itself - which is an important distinction when trying to analyze the impact of wrinkles on the identification of the action. The term “cheek rising” describes the action but imprecisely; it refers to the larger part of the face (and not exclusively to the activity of orbitalis oculi muscle). The evaluation of AU6 focused on the cheek raising and changes in the lower face, but did not empathize the role of crows-feet building around the eyes. These difficulties were overwhelmed in the second study where we evaluated the full face expressions. Regarding the effect of wrinkles on AU12 - this action was highly identified also without wrinkles and with low intensity. AU12 received very high recognition scores in all the conditions (from 4.02 to 4.33; see Table II). Thus, despite numerically higher mean scores it would be unlikely to observe the significant identification improvement for AU12 with wrinkles.

Regarding the naturalness and realism, animations of single action units were perceived as more natural and realistic compared to static ones (images) (i.e., hypothesis H1.2.b). Unexpectedly, the utilization of wrinkles was not influencing the perception of realism and naturalness (hypothesis H1.3.b). This is probably due to the simplicity of our algorithm used to generate expressive wrinkles. It also shows that even a very subtle animation (such as frowning) does not need to be highly realistic nor natural to be identified correctly.

Additionally, naturalness and realism were highly correlated. This can be explained by the fact that only very short animations were shown and the evaluation of the naturalness of the behavior was probably reduced to the quality of the displayed expressions. Finally, we also observed differences between the two VCs used in the experiment: the female character was perceived as more natural. However, this result should not be used to drive any conclusion about the role of gender as we have used only one female and only one male character in this evaluation. It may indicate, however, that the animation quality of these two characters is not comparable and that animations of one of the characters are better than the other. This leaded us to the decision to only use the female character for the further investigations.

To conclude, two factors, intensity and presentation mode, influenced the identification and perceived animation quality of the synthesized single action units. In the case of wrinkles, the results are ambiguous. Although wrinkles influence the identification of AU4, this effect was not observed for the other two AUs. Taking into account the challenges related to the experiment procedure (i.e., imprecise labels), the current results do not allow us to give a clear answer about the role of wrinkles in the perception of VC expressions; therefore, a further study was needed.

5. STUDY 2: PERCEPTION OF LAUGHTER EXPRESSIONS

Drawing on the conclusions from Study 1, we proposed the second study to further test the decoding of full face expressions. In this study the participants were shown a set of the synthesized full face expressions generated with the approach presented in Section 3.2 and asked to evaluate them with a set of criteria measuring the animation quality and its meaning. Similarly to the previous study we still focused on the three action units of laughter, i.e. AU4, AU6 and AU12, but we analyzed their perception in the context of full face expressions of laughter. Consequently we considered all plausible combinations of AU4, AU6 and AU12 in laughter expressions (see Section 2.3). While AU12 is an indispensable component of all laughter displays, the other two, namely AU4 and AU6, are optional and their presence can be manipulated.

In this study, we used two out of three factors considered in the previous study: i.e., intensity and wrinkles. The appearance of "laughter" action units i.e. AU4, AU6, and AU12 is reinforced by the use of the wrinkles. The animations display three intensity degrees (low, medium, high). The third variable of the previous study i.e. the presentation mode was not used in this study, as there exist solid findings on enhanced performance in dynamic stimuli (see Section 2.2) that were confirmed in our first evaluation study (see previous section). Finally, as results of Study 1 indicated the effect for character version, in the second study we only used the character whose animations were previously perceived more natural.

5.1 Experiment Setup

5.1.1 *Stimuli selection.* Six laughter episodes were chosen from the AVLC corpus. All the episodes were generated from the data of the same female subject (number 5) and represents three different intensity levels. For this purpose intensity annotations of AVLC corpus were used (see [Niewiadomski et al. 2012b] and Section 3.1 for details) and two low intensity, two medium intensity and two high intensity laughs were chosen.

From the original laughs, 36 different animations were generated, grouped into two experimental conditions: with wrinkles (WR) and no wrinkle (NW) (between-subject variable). To ensure that the final results do not depend only on one action unit (or one wrinkle) that is always present in the laughter displays (i.e. AU12), we applied the algorithm presented in Section 3.2 to enhance visibility of

a b c

Fig. 4. One frame of the animation in a set: A1 (Fig. 4a), A2 (Fig. 4b) and A3 (Fig. 4c).

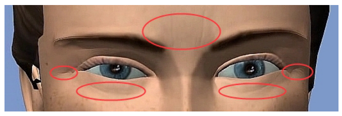

Fig. 5. Wrinkles associated with AU4 and AU6 (a close-up of the Figure 4c).

other action units. Consequently in each condition, the six episodes were displayed in three variations (within-subject variable): A1) the animation generated by the data-driven approach, A2) the animation A1 with AU6 enhanced using AUFilters, A3) the animation A1 with AU6 and A4 enhanced using AUFilters. The difference between the sets A1-A3 can be seen on Figures 4 and 5. All animations were generated with the female VC. All animations used the original audio file (laughter vocalizations).

5.1.2 *Evaluation questionnaire and hypotheses.* Participants were asked to rate the naturalness, degree of amusement and intensity of presented laughs. Importantly, participants were explicitly instructed to focus on facial expressions. All items were answered on a five point Likert scale ranging from 1 = "not at all" to 5 = "totally". When referring to the virtual character, the term "avatar" was utilized, as it is more easily understood by naive participants. In more details we asked the following questions: Q1) "How intense is the laugh of the avatar?", Q2) "How amused is the avatar?", "Q3) How natural is the laughter facial expression?". The questions Q1 and Q2 focus on the interpretation of facial expressions, while the question Q3 evaluates the animation quality. Differently to the first evaluation we used only the variable "naturalness" as the second one, i.e. "realism", was found to be highly correlated with the "naturalness".

It is important to notice that while laughter can have many meanings we focused here on the most obvious one i.e. sign of amusement. To evaluate the perceived meaning of the expression we asked to estimate its amusement level. Thus, for us, the difference in the perceived amusement shows that the message is decoded differently. Consequently, the considered factor (i.e. wrinkles) is significant for the meaning decoding and, thus it influences the communication of the intention.

Our hypotheses were:

—H2.1.a. Intensity of an original laughter video influences the perceived intensity of its animation;

—H2.1.b. Higher intensity expressions are perceived as less natural;

—H2.2.a. Presence of wrinkles influences the perceived amusement laughter expressions;

—H2.2.b. Synthesized facial wrinkles increase the perceived naturalness of laughter expressions.

Table III. Evaluation of the full face expressions of laughter.

	wrinkles								no wrinkles							
	low		medium		high		total		low		medium		high		total	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Q1	2.23	0.43	2.73	0.51	3.91	0.63	2.96	0.44	1.99	0.51	2.37	0.8	3.52	0.77	2.63	0.55
Q2	2.37	0.64	3.04	0.64	3.95	0.49	3.12	0.51	2.2	0.66	2.5	0.77	3.41	0.87	2.71	0.62
Q3	2.43	0.75	2.99	0.62	3.47	0.48	2.96	0.48	2.49	0.92	2.79	0.91	3.22	1.09	2.83	0.84

Regarding the hypothesis H.2.1.a we expected that degrees of intensity observed in the original videos of AVLC will be correctly decoded from animations. In that case, we additionally aimed to check (hypothesis H2.1.b) if the conclusion of the previous study saying that "the more intense expressions are perceived as less natural" was true also for full face expressions.

Regarding the hypotheses H.2.2.a - H.2.2.b we believed that wrinkles alter the perception of the facial expressions, by making all laughter facial actions more visible (i.e. AU12, AU6 and AU4) and thus influence the perceived meaning of the generated animation. The wrinkles can be useful to stress these single actions, helping the human observer to interpret the meaning of the whole facial behavior. We also expected that the synthesized facial wrinkles may improve the perception of naturalness.

5.2 Procedure

Participants were given a link to an online survey where they received the general instructions of the study. The instruction included that participants would be presented with animations of a laughing avatar. Participants were made aware that some of the animations might look similar, as the animations were chosen from a big pool of laughter examples and were randomly assigned to them. They were asked to answer all questions spontaneously and not to omit any answers. Participants were told that the presented stimuli contain audio, but that they should focus on the facial features.

The evaluation was organized on a set of web pages. Each web page displayed one animation episode. Participants could see the animations only once and they had to answer all the questions before seeing the next episode. Each participant evaluated only one condition (WR/NW) i.e., 18 animations. Intentionally we used between-subjects design to evaluate the wrinkles impact; otherwise the participants could easily guess the cue that is manipulated (i.e. presence of wrinkles) and bias the results.

5.3 Results

The sample consisted of 46 adult participants from 14 countries (28 females, age 23-60; mean age 30, SD = 7, most frequent countries of origin: Switzerland 24%, Poland 11%, Italy 11% and Germany 11%). Participants were randomly assigned to one of the two conditions (animations with wrinkles (WR) $n = 23$ participants; and animations without wrinkles (NW) $n = 23$ participants). First, descriptive statistics of the evaluation ratings were computed for the two experimental conditions separately. The results are given in Table III.

To test our hypotheses we ran a mixed-design MANOVA with two within-subject measures: Intensity (low, medium, high) and animation Version (A1, A2, A3), and one between-subjects variable Wrinkles (yes vs. no) and questions Q1 - Q3 as dependent variables. MANOVA showed the combined effect of Wrinkles ($F(3, 42) = 3.203$, $p < .05$; Wilk's $\lambda = .814$), and of Intensity ($F(6, 39) = 46.327$, $p < .001$; Wilk's $\lambda = .123$), but not of Version ($F(6, 39) = .668$, $p = .676$; Wilk's $\lambda = .907$).
Within-subject analysis showed a significant effect of Intensity ($F(6, 172) = 35.312$, $p < .001$; Wilk's $\lambda = .201$). In more details, the results show a significant effect of the animation intensity on the perceived intensity (question Q1; $F(2, 88) = 143.201$; $p < .001$), amusement (question Q2, $F(2, 88) = 100.775$; $p < .001$) and naturalness (question Q3; $F(1.592, 86.144) = 30.774$; $p < .001$ after application of Greenhouse-Geisser correction of sphericity ($\epsilon = .796$)). Additionally, within-subjects trends analysis

showed that perceived intensity did increase linearly to the animation intensity ($F(1,44) = 233.731$, $p < .001$). Furthermore, there was also a quadratic trend ($F(1,44) = 21.553$, $p < .001$).
Between subject analysis showed a significant effect of Wrinkles on the perceived intensity (question Q1; $F(1,44) = 5.25$; $p < .05$), amusement (question Q2; $F(1,44) = 6.213$; $p < .05$) but not on naturalness (question Q3; $F(1,44) = .416$; $p = .522$).
Finally, all interaction effects were not significant: in particular Version × Wrinkles ($F(6,39) = 2.3$, $p = .06$; Wilk‘s $\lambda = .739$), Intensity × Wrinkles ($F(6,39) = .783$, $p = .589$; Wilk‘s $\lambda = .893$), and Intensity × Version ($F(12,33) = 1.341$, $p = .243$; Wilk‘s $\lambda = .672$).

5.4 Discussion

In this evaluation, the perception of synthesized expressions of laughter was investigated. Two out of four hypotheses were confirmed. Consistently with expectations, the higher intensity animations were perceived as intense compared to the low intensity animations (hypothesis H2.1.a). Regarding the hypothesis H2.1.b the animations of intense full face laughter expressions were perceived as more natural than the low(er) intense ones. This result is somehow contradicting the results of the first evaluation: the opposite relation was observed for animations of single action units. A possible explanation can be that participants evaluated intense expressions of single actions as less natural because they were not used to this kind of expressions alienated from a context. Expressions of intense single action units (e.g., inner raise eyebrow) do not occur often in real-life interactions if they are not accompanied, proceeded or succeeded by other facial actions. Low intense single facial actions, e.g. a slight frown, might, however, may appear more often (e.g. as a sign of concentration). Consequently, intense single actions were considered less natural.

Regarding the hypothesis H2.2.a, wrinkles were shown to be important component of facial displays that may alter the interpretation of the expressions. Indeed wrinkle-based (WR) expressions were perceived as more amusing than their no wrinkles (NW) correspondences. Moreover, the effect of wrinkles was consistent and it did not depend on a particular wrinkle or action (no interaction between Wrinkles and Version). For more generalized conclusions, this study should be replicated with other expressions than laughter.

It is also important to notice that the animations with wrinkles were considered more amusing and intense compared to the animations without wrinkles. This is probably due to AU6 that is more visible with wrinkles. If so, it would confirm our explanation of the results of the first study regarding this particular action. Indeed the results of the second study suggest that the wrinkles of AU6 are noticed and decoded correctly. Thus the hypothesis about the miscomprehension of the label used in the first study seems to be a plausible explanation of contradicting results of these two experiments. Consequently, we may assume that wrinkles influence the perceived meaning of full face expressions, because they influence the perception of single action units locally.

Differently to our expectations (but accordingly to the results of the first study) wrinkles did not improve the perceived quality of animation (hypothesis H2.2.b). This effect can be explained by the simple rendering method used to generate wrinkles. More advanced but also more computationally expensive methods might give different results (see, for example, [Courgeon et al. 2009a].

6. GENERAL DISCUSSION

In this paper we use virtual characters to analyze factors that influence the perception of synthesized expressions. Our research was driven from two important facts:

—Firstly, since the past few years, there is a strong increase of interest in VC technology, in particular regarding applications of serious games and personal assistants. For such applications, it is crucial

Table IV. The results of this study in the comparison of the previous studies on the perception of the human and synthesized expressions. Importantly in this work, we use different set of expressions.

Literature	Intensity	Presentation	Wrinkles
Perception of real (human) expressions	Previous results confirmed	Previous results confirmed	no previous works
Perception of synthesized expressions	Previous results confirmed	**New results showing a significant role of the factor**	**New results showing a significant role of the factor**

that the VC is capable of displaying accurately facial expressions of emotions. Our work brings new knowledge on how synthesized expressions are perceived and which cues are important in the perception of expressions.

—Secondly, using virtual characters in studies on the perception of facial expressions allows an experimenter to control any small detail of the facial expressions. This is not always possible to do so with human's expressions, as one cannot easily control certain expressive features such as the appearance of wrinkles.

Below we present our concluding remarks regarding both of these aspects.

Presentation Mode. In existing studies on VCs expression, the role of the presentation mode in the perception of synthesized facial expressions was ambiguous. Moreover, most of these studies were limited only to the expressions of basic emotions. While some studies showed the presentation mode is important for the recognition of anger and fear but not for the recognition of the other four basic expressions ([Katsyri and Sams 2008]), other works (e.g. [Noël et al. 2006]) did not reveal an effect of the presentation mode. In our own study, we observed an effect of the presentation mode: the synthesized facial expressions were more easily identified from animations than from images.
Our result is also consistent with the results of studies on the perception of human expressions (e.g. [Ambadar et al. 2005], see Section 2.2). The reason why our results differ from some of the previous results on the perception of synthesized expressions is probably related to the choice of the stimuli: we use more subtle expressions. In consequence, our study confirms that the presentation mode is important in the perception of non basic synthesized expressions.

Intensity. In concordance with our hypotheses we observed that the intensity of synthesized expressions influence their identification rate. This is in agreement with previous works on virtual characters (e.g. [Bartneck and Reichenbach 2005]).
Regarding the intensity factor in the human full face expressions, Hess et al. [Hess et al. 1997] showed that, for basic emotions, perceived intensity was linear to the physical intensity. Focusing more particularly on expressions of positive affects, the expression of happiness is not equal to the expression of hilarious laughter. However, both expressions contain the same action units: namely, AU6 and AU12. As in Hess et al. ([Hess et al. 1997]), we found that the perceived intensity of the expression of laughter is proportional to the intensity of the corresponding synthesized facial expression.

Wrinkles. To our knowledge, studies on the perception of human facial expressions analyzed only age-related wrinkles. So far, no experimental studies focusing on the role of expressive (i.e., AU-activity related) wrinkles have been performed. This lack of studies could be due the fact that such wrinkles are not easily manipulated and controlled on human faces. Thus, it is difficult to create stimuli to study. Using VC allows a precise control over facial features and expressions and offers a mean to create precisely manipulated stimuli. Our results show that wrinkles influence the identification and the perception of the meaning of expressions. While the age-related wrinkles in humans seem to decrease the clarity of the emotional message (see Section 2.2) the expressive wrinkles in VCs, on the contrary, seem to help the observer to decode the expression and its meaning. The wrinkles allow one to identify correctly an action. Consequently, they change or reinforce the perceived meaning of the full face expression. It is important to notice that the algorithm used to generate wrinkles was rather simple; it did not influence the geometry of the virtual model but it applied the wrinkles on the surface. We believe this explains why the perceived animation quality (its naturalness and realism) was not influenced by adding the wrinkles.

More globally, this work, in contrast with previous studies, focuses on non basic expressions, and it brings new results regarding their identification and perception. Our results show a significant role of presentation mode. Also the presence of wrinkles is important in identifying and perceiving the facial expressions of non basic emotions.

7. CONCLUSION

This study investigated the factors influencing the decoding of isolated action units, as well as of full face facial expressions in the context free displays of laughter. First we studied factors that may influence the identification of single actions (e.g., frown or cheek raising), namely their intensity, their presentation mode, and the application of wrinkles. Results of our first study showed that single facial actions are better identified when they are dynamic and of higher intensity. At the same time intense expressions of single facial actions were perceived less natural and less realistic. Also, wrinkles improve the identification of an action significantly. Nevertheless, it is very seldom that action units appear in isolation in interaction. Thus, additionally the perception of full face expressions was analyzed in the second study. It was shown that wrinkles impact the meaning attached to facial expressions. In particular, they increased the perceived amusement of the laughter expressions. Thus, all three factors (wrinkles, intensity, presentation mode) influence the identification and meaning of a synthesized expression, while two of them (intensity and presentation mode) impact the perceived animation quality.

It is important to notice that several other variables may influence the perception of expression such as gender or age of the decoder. They should be considered in future works. Secondly in our studies we have used intra-culture design although we are aware that expressions might be decoded and evaluated differently in different cultural contexts (see e.g. [Elfenbein and Ambady 2002]). The other limitation of our studies is related to rather simple rendering technique applied to generate animations. However, the significant results obtained in these studies makes us believe that the role of wrinkles might be even more significant if more expensive computationally techniques were used. Additionally, the second study is limited to the data of one subject. A more extended dataset should be used in future works. Although this work was dedicated to the laughter expressions, the three analyzed action units are components of expressions of many other communicative intentions. Thus we expect that the role of wrinkles, presentation mode and intensity goes beyond the displays of laughter.

We can draw several conclusions related to the results of our two studies. Firstly, the role of intensity and presentation mode in the perception of VC expressions was found to be concordant with our expectations. These two factors have similar impact on the perception of human and virtual character expressions. Secondly, expressive wrinkles in the synthesized expressions have an important role and, consequently, they should not be neglected when creating virtual characters able to communicate emotions nonverbally. VC will be more and more present in our every day life, and the way they express emotions will have an impact on humans, i.e. the addressees of this technology. In past, it was shown that several appearance related factors have an impact on the perception of a VC e.g., its gender [Ochs and Pelachaud 2013], or race [Rossen et al. 2008]. According to our results, wrinkles also influence the identification and meaning of synthesized expressions. Thus, when designing virtual characters, the specification of wrinkles should be considered as expressive wrinkles bring important information to the observer, allowing him to decode the meaning of the emotional message. Expressive wrinkles impact the perceived meaning of full face expressions, because they influence the perception of single action units locally. Finally, the results of our studies show also that virtual characters are a valid tool to study various meanings of subtle emotion expressions (as manipulation of only small components of a display is perceivable by humans who may even associate these changes to particular meanings. Our VC able to display expressive wrinkles was recently used by Hyniewska in her study of the impact of nonverbal cues on the perception of internal states of the displayer [Hyniewska 2013].

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n 270780 ILHAIRE. Furthermore, the authors would like to thank Bing Qu, Satish Pammi, and Jing Huang (Telecom Paristech) for helping in creation of the animations used in the studies presented in this paper.

REFERENCES

AMBADAR, Z., SCHOOLER, J. W., AND COHN, J. F. 2005. Deciphering the enigmatic face. The importance of facial dynamics in interpreting subtle facial expressions. *Psychologial Science 16,* 5, 403–410.

AYLETT, R., PAIVA, A., DIAS, J., HALL, L., AND WOODS, S. 2009. Affective agents for education against bullying. In *Affective Information Processing*, J. Tao and T. Tan, Eds. Springer London, 75–90.

BARTNECK, C. AND REICHENBACH, J. 2005. Subtle emotional expressions of synthetic characters. *International Journal Human-Computer Studies 62,* 3, 179–192.

BASSILI, J. N. 1979. 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, 2049–2058.

BLINN, J. 1978. Simulation of wrinkled surfaces. In *Proceedings of SIGGRAPH '78*. ACM, NY, USA, 286–292.

COURGEON, M., BUISINE, S., AND MARTIN, J.-C. 2009a. Impact of expressive wrinkles on perception of a virtual characters facial expressions of emotions. In *Intelligent Virtual Agents*, Z. Ruttkay, M. Kipp, A. Nijholt, and H. Vilhjlmsson, Eds. Lecture Notes in Computer Science Series, vol. 5773. Springer Berlin Heidelberg, 201–214.

COURGEON, M., CLAVEL, C., AND MARTIN, J.-C. 2009b. Appraising emotional events during a real-time interactive game. In *Proceedings of the International Workshop on Affective-Aware Virtual Agents and Social Robots (AFFINE '09)*, G. Castellano, J.-C. Martin, J. Murray, K. Karpouzis, and C. Peters, Eds. ACM, New York, NY, USA.

DARWIN, C. 1872. *The expression of emotion in man and animal*. John Murray, London.

DRACK, P., HUBER, T., AND RUCH, W. 2009. The apex of happy laughter: A FACS-study with actors. In *Current and Future Perspectives in Facial Expression Research: Topics and Methodical Questions*, E. Banninger-Huber and D. Peham, Eds. Universitat Innsbruck, 32–37.

EKMAN, P. AND FRIESEN, W. V. 1975. *Unmasking the Face. A guide to recognizing emotions from facial clues*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

EKMAN, P., FRIESEN, W. V., AND HAGER, J. C. 2002. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto.

ELFENBEIN, H. AND AMBADY, N. 2002. On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin 128*, 203–235.

FRANK, M. G. AND EKMAN, P. 1993. Not all smiles are created equal: The differences between enjoyment and non-enjoyment smiles. *Humor: International Journal of Humor Research 6*, 9–26.

HESS, U., ADAMS, R. B., SIMARD, A., STEVENSON, M. T., AND KLECK, R. E. 2012. Smiling and sad wrinkles: Age-related changes in the face and the perception of emotions and intentions. *J. of Experimental Social Psychology 48,* 6, 1377 – 1380.

HESS, U., BLAIRY, S., AND KLECK, R. 1997. The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior 21,* 4, 241–257.

HYNIEWSKA, S. 2013. Non-verbal expression perception and mental state attribution by third parties. PhD thesis.

KAISER, S. AND WEHRLE, T. 2001. Appraisal processes. In *Emotion: Theory, Methods, Research*, K. Scherer, A. Schorr, and F. Johnstone, Eds. Oxford University Press, 285–300.

KATSYRI, J. AND SAMS, M. 2008. The effect of dynamics on identifying basic emotions from synthetic and natural faces. *International Journal Human-Computer Studies 66,* 4, 233–242.

KELTNER, D. 1995. Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology 68*, 441–454.

KRUMHUBER, E., MANSTEAD, A. S., COSKER, D., MARSHALL, D., AND ROSIN, P. L. 2009. Effects of dynamic attributes of smiles in human and synthetic faces: A simulated job interview setting. *Journal of Nonverbal Behavior 33,* 1, 1–15.

MARTIN, R. A. AND KUIPER, N. A. 2009. Daily occurrence of laughter: Relationships with age, gender, and type A personality. *International Journal of Humor Research 12,* 4, 355–384.

MELO, C., CARNEVALE, P., AND GRATCH, J. 2012. The effect of virtual agents emotion displays and appraisals on peoples decision making in negotiation. In *Intelligent Virtual Agents*, Y. Nakano, M. Neff, A. Paiva, and M. Walker, Eds. Lecture Notes in Computer Science Series, vol. 7502. Springer Berlin Heidelberg, 53–66.

NIEWIADOMSKI, R., BEVACQUA, E., LE, Q. A., OBAID, M., LOOSER, J., AND PELACHAUD, C. 2011. Cross-media agent platform. In *Proceedings of the 16th International Conference on 3D Web Technology*. ACM, Paris, France, 11–19.

NIEWIADOMSKI, R., HUANG, J., AND PELACHAUD, C. 2012a. Effect of facial cues on identification. In *Proceedings of the 25th Annual Conference on Computer Animation and Social Agents (CASA 2012)*. Singapore, 37–44.

NIEWIADOMSKI, R., URBAIN, J., PELACHAUD, C., AND DUTOIT, T. 2012b. Finding out the audio and visual features that influence the perception of laughter intensity and differ in inhalation and exhalation phases. In *Proceedings of 4th International Workshop on Corpora for Research on Emotion, Sentiment & Social Signals, LREC 2012*. Istanbul, Turkey.

NOËL, S., DUMOULIN, WHALEN, T., AND STEWART, J. 2006. Recognizing emotions on static and animated avatar faces. In *IEEE International Workshop on Haptic Audio Visual Environments and their Applications*. Ottawa, Canada, 99–104.

OCHS, M., NIEWIADOMSKI, R., BRUNET, P., AND PELACHAUD, C. 2012. Smiling virtual agent in social context. *Cognitive Processing 13,* 2, 519–532.

OCHS, M. AND PELACHAUD, C. 2013. Socially aware virtual characters: The social signal of smiles. *IEEE Signal Processing Magazine 30,* 2, 128–132.

OSTERMANN, J. 2002. Face animation in MPEG-4. In *MPEG-4 Facial Animation - The Standard Implementation and Applications*, I. Pandzic and R. Forchheimer, Eds. Wiley, England, 17–55.

QU, B., PAMMI, S., NIEWIADOMSKI, R., AND CHOLLET, G. 2012. Estimation of FAPs and intensities of AUs based on real-time face tracking. In *FAA The 3rd International Symposium on Facial Analysis and Animation*. Vienna, Austria.

ROSSEN, B., JOHNSEN, K., DELADISMA, A., LIND, S., AND LOK, B. 2008. Virtual humans elicit skin-tone bias consistent with real-world skin-tone biases. In *Intelligent Virtual Agents*, H. Prendinger, J. Lester, and M. Ishizuka, Eds. Lecture Notes in Computer Science Series, vol. 5208. Springer Berlin Heidelberg, 237–244.

RUCH, W. 1993. Exhilaration and humor. In *The handbook of emotions*, M. Lewis and J. M. Haviland, Eds. Guilford, New York, 605–616.

RUCH, W. AND EKMAN, P. 2001. The expressive pattern of laughter. In *Emotion qualia, and consciousness*, A. Kaszniak, Ed. Word Scientific Publisher, 426–443.

RUCH, W. F., HOFMANN, J., AND PLATT, T. 2013. Investigating facial features of four types of laughter in historic illustrations. *European Journal of Humor Research 1*, 98–118.

RYMARCZYK, K., BIELE, C., GRABOWSKA, A., AND MAJCZYNSKI, H. 2011. EMG activity in response to static and dynamic facial expressions. *International Journal of Psychophysiology 79,* 2, 330 – 333.

SARAGIH, J. M., LUCEY, S., AND COHN, J. F. 2011. Deformable model fitting by regularized landmark mean-shift. *International Journal of Computer Vision 91*, 200–215.

SATO, W. AND YOSHIKAWA, S. 2004. Brief report the dynamic aspects of emotional facial expressions. *Cognition & Emotion 18,* 5, 701–710.

SCHERER, K. R. 2001. Appraisal considered as a process of multilevel sequential checking. In *Appraisal Processes in Emotion: Theory, Methods, Research*, K. Scherer, A. Schorr, and T. Johnstone, Eds. Oxford University Press, 92–119.

SCHERER, K. R., WRANIK, T., SANGSUE, J., TRAN, V., AND SCHERER, U. 2004. Emotions in everyday life: Probability of occurrence, risk factors, appraisal and reaction patterns. *Social Science Information 43*, 499–570.

URBAIN, J., BEVACQUA, E., DUTOIT, T., MOINET, A., NIEWIADOMSKI, R., PELACHAUD, C., PICART, B., TILMANNE, J., AND WAGNER, J. 2010. The AVLaughterCycle Database. In *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*. European Language Resources Association (ELRA), Valletta, Malta, 2996–3001.

WEYERS, P., MÜHLBERGER, A., HEFELE, C., AND PAULI, P. 2006. Electromyographic responses to static and dynamic avatar emotional facial expressions. *Interantional Journal of Psychophysiology 43,* 1, 450–453.

YOUNG, A. W., ROWLAND, D., CALDER, A. J., ETCOFF, N. L., SETH, A., AND PERRETT, D. I. 1997. Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition 63,* 3, 271 – 313.