FACSHuman, a software program for creating experimental material by modeling 3D facial expressions

Michaël Gilbert[1] · Samuel Demarchi[1] · Isabel Urdapilleta[1]

Abstract

This paper presents the FACSHuman software program, a tool for creating facial expression materials (pictures and videos) based on the Facial Action Coding System (FACS) developed by Ekman et al. (2002). FACSHuman allows almost all the Action Units (AUs) described in the FACS Manual to be manipulated through a three-dimensional modeling software interface. Four experiments were conducted to evaluate facial expressions of emotion generated by the software and their theoretical efficiency regarding the FACS. The first study (a categorization task of facial emotions such as happiness, anger, etc.) showed that 85% of generated pictures of emotional expressions were correctly categorized. The second study showed that only 82% of the most-used AUs were correctly matched. In the third experiment, two independent FACS coders rated 47 AUs generated by FACSHuman using the standard methodology used in this kind of task (AU identification). Results showed good-to-excellent codification rates (64% and 85%). In the last experiment, 54 combinations of AU were evaluated by the same FACS coders. Results showed good-to-excellent codification rates (68–82%). Results suggested that FACSHuman could be used as experimental material for research into nonverbal communication and emotional expression.

Keywords Facial expressions · Nonverbal communication · FACS · Emotion · Avatars

Background

Research into non-verbal communication and particularly facial expression has often required the use of specific experimental material related to the body or the face. The creation of this material can be produced by different methods depending on the experimental goals. This can be a databank of pictures or videos selected for emotional induction (Bänziger et al., 2009). For facial expressions, it can be produced by actors directed by the researcher (Aneja et al., 2017; Happy et al., 2017; Langner et al., 2010; Lucey et al., 2010; Mavadati et al., 2013; Mollahosseini et al., 2017; Sneddon et al., 2012; Valstar & Pantic, 2010) or by recording the spontaneous reactions of subjects exposed to selected stimuli to induce targeted facial expressions (Tcherkassof et al., 2013). This material is mainly composed of pictures and videos showing postures of the body and face.

The production of a software program to model facial expressions and fine movements of the face would allow researchers to create experimental material based on the characteristics of the Facial Action Coding System (FACS) (Ekman et al., 2002). These pictures seem to reflect a possible social interaction or basic emotion depending on the researcher's theoretical position (Crivelli et al., 2016; Crivelli & Fridlund, 2018).

More recently, several software programs have been produced and released to researchers (Amini et al., 2015; Villagrasa & Susín Sánchez, 2009). These validated programs are however not completely free of rights or usable on any operating system. They are dependent on commercial 3D engines, respectively Haptek 3D-characters for the HapFACS software program (Amini et al., 2015) and FaceGen Modeller for the FACSGen (Krumhuber et al., 2012). Nevertheless, HapFACS and FACSGen software programs are free of charge and available upon request from their respective research laboratories.

✉ Samuel Demarchi
samuel.demarchi@univ-paris8.fr

[1] Department of Psychology, Université Paris 8, 2, rue de la Liberté, 93526 Saint-Denis CEDEX, France

The present paper

This paper presents a software program that allows researchers to produce the experimental material they need to create facial

expression materials (pictures and videos) based on the FACS developed by Ekman et al. (2002). FACSHuman was developed under an open-source license and is based on the characteristics of the two previously cited software programs. It extends the possibilities by allowing users to modify and redistribute it. It is based exclusively on open-source software programs (MakeHuman, Blender, Gimp) and the Python programming language usually used in academic research. Four studies were conducted to evaluate the relevance and the accuracy of this software.

FACSHuman software architecture

FACSHuman is based on the Makehuman software program and the FACS. It consists of three additional plugins for Makehuman. The main goal of FACSHuman is to produce experimental material such as images and animations, material that can meet fairly high criteria of realism, esthetics, and morphological precision. With this software program, the muscles of the face, eyes, and tongue can now be manipulated, and facial features, the color of the skin and eyes, and the skeletal structure can be customized via Makehuman.

MakeHuman software program

Makehuman was selected as a framework for the development of the software program presented in this article. It is a free software program and can be extended by external plugins. Its 3D rendering engine is based on OpenGL technology. It is used primarily for the creation of three-dimensional avatars in the video game industry and 3D recreation. This software program allows the entire human body to be modeled, including the morphology of the face and its main features as well as gender, race, and age from babies to older adults. The color of the skin and the eyes can be changed, modified or imported. The possibilities are unlimited in terms of combinations of age, gender, race. These parameters can be mixed together in the MakeHuman program, allowing the users to create as many character identities as they need. These characters can be saved and shared and reused for other experiments or exported into a third 3D software program such as Blender or Unity 3D. Previously cited software programs are limited to the face or the head and the bust. MakeHuman, which we used as a framework for our plugins, allows the creation of body shape and postures.

The Facial Action Coding System

The FACS was developed by Paul Ekman (Ekman et al., 2002). This system is mainly used in research in nonverbal human activities and behaviors but also in research into the creation of avatars and artificial agents. It allows the fine movements of the face to be coded and facilitates the dissemination of facial configurations observed or used in research. This system breaks the face down into Action Units (AUs), which correspond to the muscular movements of the face observable in humans (as shown in Fig. 1). During the observation, the coding of each AU is characterized by two pieces of information (number of Action Units and intensity). For instance, 1B + 2C + 5D is a code of an observed facial configuration. Table 1 contains some examples of AUs taken from the learning manual of this coding method (Ekman et al., 2002).

This coding system requires long training estimated at more than 150 h (Ekman et al., 2002). A certification is available and validates the acquired skills. This system of codification of facial expressions was selected for the development of this software program.

Creation process of textures in 3D model used in the additional plugins

The 3D animation of both the human body and face can be created by various technical means such as the mixing of structures, animation by bones, manipulation of textures (Magnenat-Thalmann & Thalmann, 2004) and muscular geometric modeling (Kähler et al., 2001). The animation technique selected for the development of this software program was *BlendShapes* or *morphing* of target expressions (mixture of structures). All targets used in FACSHuman plugins were created with the Blender software and imported into Makehuman. These different targets are available and can be easily and simply manipulated within the FACSHuman plugin.

Software development presentation

A set of three additional plugins to the MakeHuman software has been developed for facial expression creation (plugin 1), facial animation (plugin 2), and for scene editing (plugin 3). These allow the user to produce images of a chosen quality (resolution), to make sets of still images of progressive intensity, to mix expressions between them and to generate videos.

There are several coding systems for facial movements, but the main ones are MPEG-4 facial animation (part of the Face and Body Animation proposed by Pandzic & Forchheimer, 2002) and the FACS. The work presented here is based on the FACS, which was created and revised by Ekman (Ekman et al., 2002). The FACS is used in research into nonverbal communication and emotional facial expressions as well as in artificial facial animation creation projects (Bennett & Šabanović, 2014; Dalibard et al., 2012; David et al., 2014).
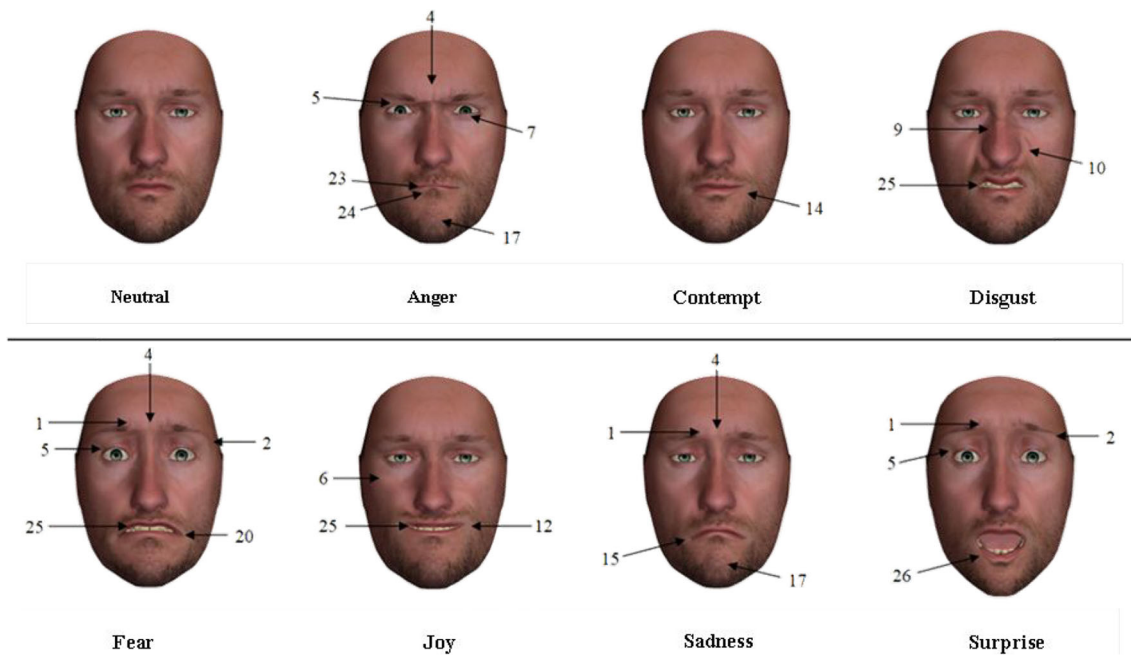
Fig. 1 Sample of coded AU with FACS

FACSHuman facial expressions creation tool (plugin 1)

The first FACSHuman plugin developed, which is presented here, allows complex facial expressions to be created, and the elements of muscular movements of the face, skin, and eyes to be defined as well as those of the jaw and the head (see Fig. 2). On the left of the screen, users can move sliders to blend predefined emotions or create their own facial expressions by increasing AU intensity one by one. On the right side of the screen, users can adjust the camera angle, its zoom function, and its position, and load and save facial expressions defined by their FACS code. Users can also program the parameters of the numbers and characteristics of the batch picture processing. In this plugin, the user can define the number of images he/she wishes to create. This allows a gradual variation to be created in the intensity of the expression, created to be used, for example in experiments on recognition thresholds (see Fig. 3).

The different AUs implemented in the software are combinable and can be mobilized on a time frame for the creation of macro and micro expressions and complex animated expressions. The animations and images created have a transparent background. The researchers thus have the freedom to present the modeled faces on a colored background or an image of their choice.

An emotional mixer is available in addition to the different manipulatable Action Units. It allows the users to create blended emotions. It uses the nine emotions described in the EmFACS (Ekman et al., 2002).

This plug-in allows users to manipulate the following modeled AUs (Appendix A1, A2, and A3). The organization of the different AU categories on the interface is based on the one used in the Score Sheet of the FACS manual.

FACSAnimation Tool (plugin 2)

The second plugin FACSAnimation tool (FANT) is a plugin for creating animations of facial expressions. The user can compose, create, and record animations by direct creation

Table 1 Sample of FACS AUs

Action units	FACS name	Muscular basis
1	Inner Brow Raiser	Frontalis, Pars Medialis
2	Outer Brow Raiser	Frontalis, Pars Lateralis
4	Brow Lowerer Depressor	Glabellae; Depressor Supercilii; Corrugator

Notes. The first column is the number assigned to the action unit, the second column is its location on the face and potential movement, and the third column is the corresponding muscle of the face. The quantification of the observed intensity is noted from an 'A' minimum observable intensity entitled 'Trace', to 'E' (maximum intensity). This quantification does not follow a linear curve (Ekman et al., 2002)
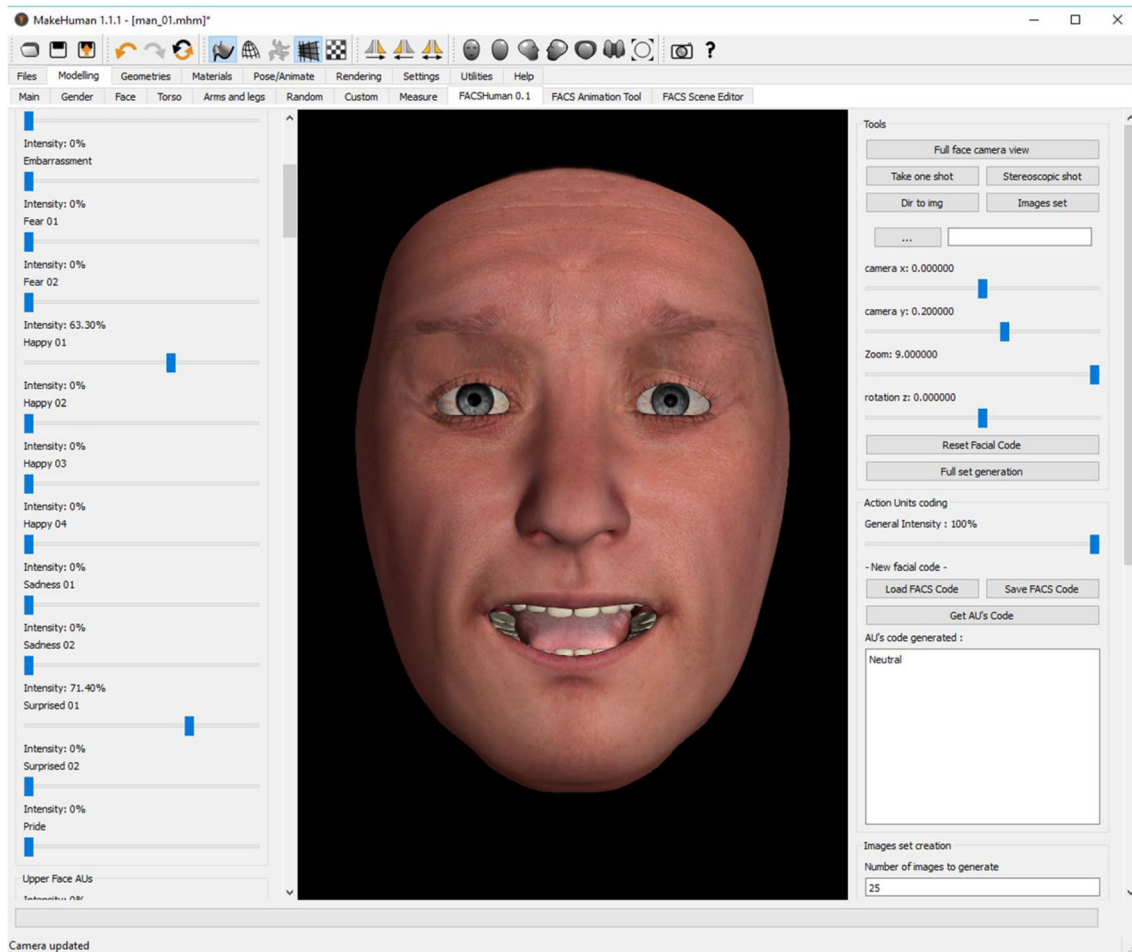
Fig. 2 FACSHuman user interface

or by mixing different expressions created in the facial expression creation tool. The batching of images is done either by modeling or by using an already recorded expressive configuration (saved in .facs files).

This image generation mode advances the intensity of each AU in a homogeneous and joint manner by following a linear progression and the maximum intensity characteristics of each AU defined by the user. The total duration of the video is controlled by the frame rate determined by the user and the number of images chosen during batch creation processing. As a result, intensity progression can be adjusted by these parameters. For the same framerate, the more images were used to create the animation, the more the video will be slowed down.

When this option is used, the progression follows the user-defined time characteristics in the FANT animation plugin. A parameterizable timeline in the FANT plugin is available and allows the creation of complex expressions such as moving from one expression to another with an apex area or using the data available for the transition between several predefined expressions. AU movements are divided into three parts. Initial intensity, apex, and final intensity. The apex areas hold the AUs at their respective maximum intensity chosen during the facial expression creation process, from the start to the stop position defined for them on the timeline. For more flexibility, apex start and stop intensity can be defined singly. In this way, users can create more complex movements (Fig. 4).

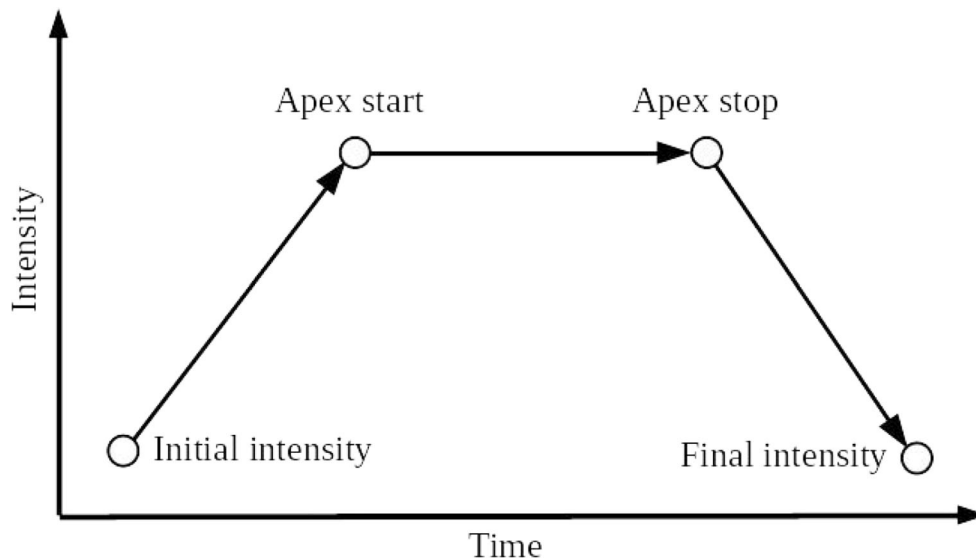**Fig. 3** Intensity progression of facial movement

Fig. 4 Characteristics of one Action Unit defined on the timeline

This also offers free creative possibilities directly in the plugin and allows the design of non-linear complex expressions that are closer to what is observable on a human face.

The timeline allows the mix of an unlimited number of expressions, each of the AUs with the characteristics (duration, start, end, intensity) of an AU described in the Investigator guide (Ekman et al., 2002). The FANT plugin allows unlimited use in numbers of an identical AU within an animation such as Expression 1 to Expression 2 to Expression 3 ... (Fig. 5). Users can choose as many AUs as needed and define the timeline characteristics for each of them inside the plugin user interface.

The target expressions are characterized by values (value from 0 to 1 of the number of images defined during the creation of the set of images). When users save their work, for each of the AUs used in the expressions, a section is recorded in a JSON file as follows:

```
"AU Name": {
 "Event_ID" {
   "image_start": "0", (Action start on the timeline)
   "image_stop": "1", (Action end on the timeline)
   "intensity_start": "0 », (Start intensity)
   "intensity_stop": "0 », (End intensity)
   "apex_start": "0.25 », (Apex start)
   "apex_stop": "0.9 », (Apex end)
   "apex_intensity_start": "1", (Apex start intensity)
   "apex_intensity_stop": "1" (Apex end intensity)
 }
}
```

The generation of synthetic animated facial expressions must be designed in a realistic way so that the stimulus presented to the subject does not provoke an effect such as the Uncanny Valley (Burleigh et al., 2013; Ferrey et al., 2015). This effect occurs by presenting a stimulus that resembles a human to a subject. It often has the effect of causing an emotional reaction of greater or lesser intensity that would partly be the result of difficulty in categorizing the object presented (Yamada et al., 2013). To minimize this effect, and to produce
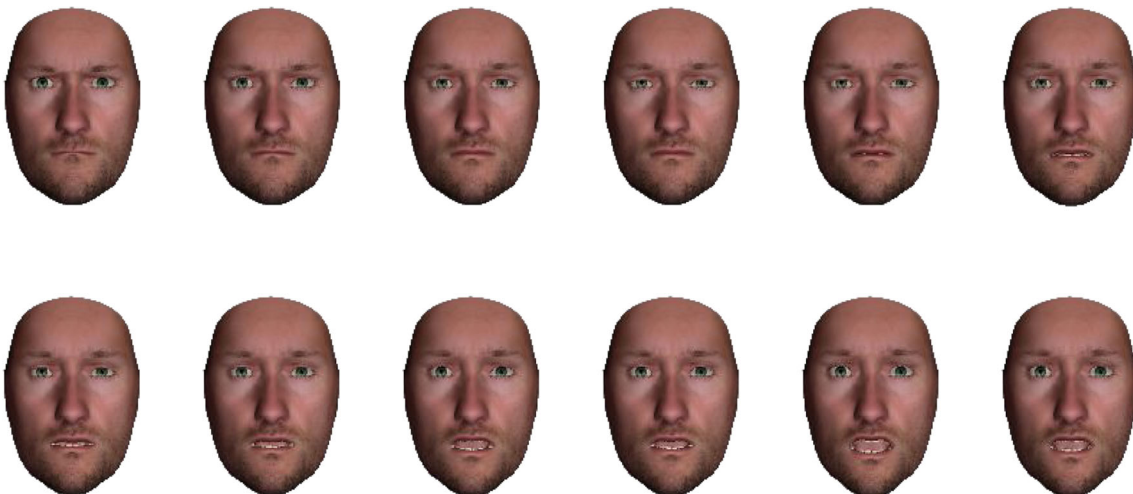

Fig. 5 Sequence produced using the FANT module, example of progression from an expression of anger to an expression of surprise

non-linear animations (Cosker et al., 2010), the FANT plugin allows users to define the temporal presentation and intensity characteristics of each AU implemented in the expression targeted by the creation of the user. For each AU, the start and stop position as well as the evolution of individual intensity can be defined. This results in observable asynchronous movements (Fig. 6). It thus allows users to transpose observed facial expressions or those taken from databases such as DynEmo (Tcherkassof et al., 2013).

FACSSceneEditor (plugin 3)

The third and last plugin FACSSceneEditor (FSCE) defines the lighting of the scene. As in a photographic studio, the user has the opportunity to place lighting around the face to create different types of staging and change their characteristics. For each added light, users can define their positions on x, y, and z axes, choose the color on a color picker, specify the light level and the specular reflection as well as other parameters available with OpenGl technology (see Fig. 7). This characteristic of the plugin results in constant lighting of the scene producing a stable environment for the generation of images, whatever the model selected.

Video production

The creation of still image batches allows the generation of videos whose number of frames per second can be defined, as well as a pause range. This pause (of a determined number of images) on an image will simulate the expressive apex when using the generation of a simple sequence (a single target

expression). Neutral to neutral image generation can be performed from a configuration file for a particular facial expression, but also by using the possibilities offered by the creation of complex animations by timeline. Users also have the ability to navigate within an animation in the software, which allows them to select a specific moment, adjust its intensity and generate a batch of images from a chosen moment.

Software portability and modifications

Add-on plugins were programmed with the same Python computer language as Makehuman. This language is widely used for academic purposes and is an interpreted language. This feature gives it great flexibility of use and the ability for users to freely modify the source code of the plugins or add more functionality if needed. Users can also automate the creation of large volumes of images or videos without external intervention. All software and plugins presented here are released under an open-source license. This software can be freely distributed, modified, and distributed in accordance with the conditions related to this type of license. This software and its plugins can be used with the main operating systems.

Experimental evaluation presentation

Four studies were conducted to carry out an evaluation of images and videos produced by the FACSHuman software program. In Experiment 1 we asked non-FACS coders to categorize emotional facial expressions created with FACSHuman and encoded from the pictures found in the
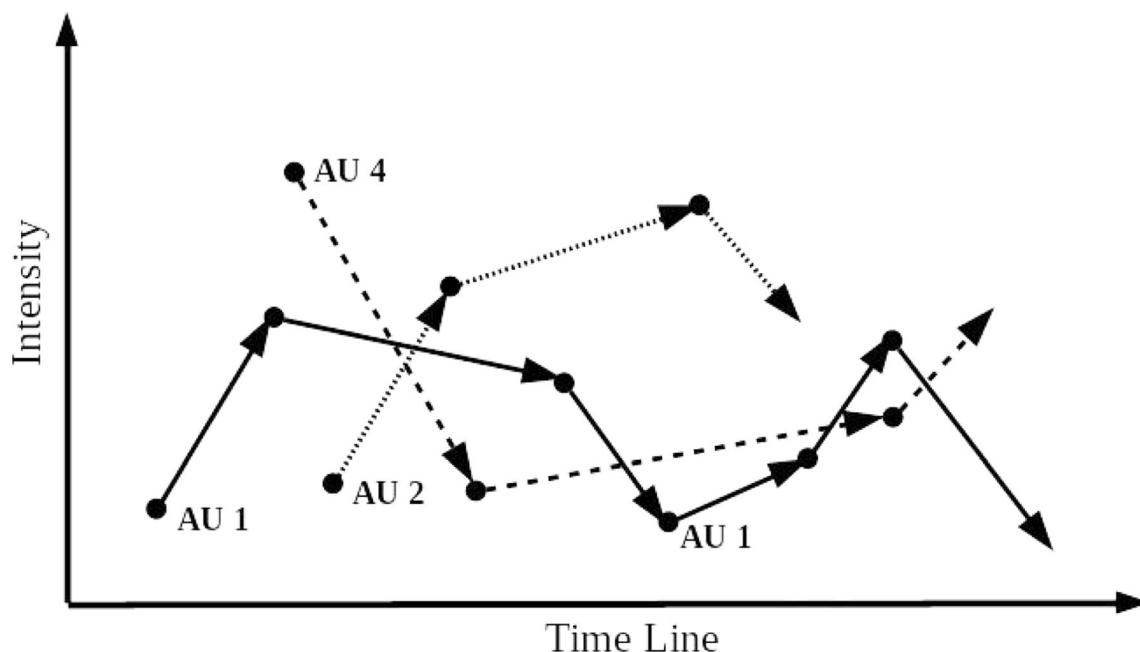

Fig. 6 Breakdown of a timeline and composition of AUs intensity

Fig. 7 Scene editor and lighting possibilities

Pictures of Facial Affect (POFA) (Ekman, 1976). In the second experiment AUs alone were evaluated by non-FACS coders in comparison to the one described in the FACS manual. In Experiments 3 and 4, the accuracy of the different facial AUs and expressions described in the FACS (Ekman et al., 2002) were evaluated by two certified FACS coders.

Experiment 1

This evaluation study of the FACSHuman software program focused on the distinction of facial configurations produced with the software and considered as representative of an emotional experience (Cigna et al., 2015; Darwin et al., 1998; Dodich et al., 2014; Ekman, 1971; Ekman, 1976; Ekman, 1992; Ekman & Oster, 1979).

Method

Participants Forty-three participants, 30 women (M_{age} = 42.1, SD_{age} = 13.44) and 13 men (M_{age} = 43.54, SD_{age} = 9.90), were engaged in this experiment via the LinkedIn professional network. They were recruited by the internal messaging system. One participant did not complete the age and gender part of the questionnaire. The participants validated, via a checkbox, a form giving free, informed and express consent before the experiment. The experiment could not begin without the validation of the form. It took place on an internet browser on the participant's computer and respected their anonymity.

Materials The experimental material consisted of 42 black-and-white images representing faces in front view for the experimental part. Six different avatars were used (three men, three women) with seven images (one neutral and six emotional facial expressions) for each avatar for a total of 42 images. Different avatars were used to avoid the reinforcement process due to multiple exposure of the same emotion with the same face. Seven pictures with the same characteristics were produced for the training part. The use of black and white was justified as an element of comparison with existing databases and in particular the POFA (Ekman, 1976) as well as images from the FACS manual (Ekman et al., 2002). Previous studies have shown that there is no significant difference in recognition or categorization tasks between color and black-and-white images (Amini et al., 2015; Krumhuber et al., 2012). These faces were produced with GIMP and Blender software programs for the texture and the FACSHuman plugin for the modeling of facial expressions.

The construction of the gender and the criteria used for the creation of the images were based on characteristics described in previous studies like facial features such as the eye region, eyebrow shape, skin texture and chin shape (Baudouin & Humphreys, 2006; Bruce et al., 1993; Bruyer et al., 1993). We used six facial expressions of basic emotions (anger, disgust, fear, happiness, sadness, surprise) and a neutral expression for each modeled face. This material was produced according to facial expression configurations from the best evaluated pictures found in the evaluation table provided with the POFA database. All these facial expressions were coded by a certified FACS coder. Table 2 presents the detailed FACS codes.

The experiment was constructed using the JSPsych (de Leeuw, 2015), which allows the creation of experiments that run in an internet browser. This feature offers the possibility of organizing experimental tests on the participant's computer with response times equivalent to those observed in the laboratory (de Leeuw, 2015; de Leeuw & Motz, 2016; Pinet et al., 2017; Reimers & Stewart, 2015). The experiments can also be used in the laboratory. Data were collected anonymously and stored in a MySQL database.

Procedure Participants were asked to categorize FACSHuman generated pictures that were presented according to the six basic emotions (anger, disgust, fear, happiness, sadness, surprise) and neutral.

The experiment began with a training phase on the use of the interface. Buttons at the bottom of the screen were used to navigate from one screen to another. The participants

Table 2 FACS codes used to create facial expressions

Expression	FACS codes
Anger	4 + 5 + 7 + 23 + 25 + 26
Disgust	6 + 7 + 9 + 10 + 17
Fear	1 + 2 + 4 + 5 + 20 + 25 + 26 + 58
Surprise	1 + 2 + 5 + 25 + 26 + 54 + 63
Sadness	1 + 4 + 7 + 17 + 20 + 23 + 43 + 52 + 54 + 62 + 64
Happiness	6 + 7 + 12

interacted with the interface using the pointing device available on the computer. During the training and experimentation phase, participants were instructed to answer the question: 'What emotion is this person expressing?'. They responded by clicking several buttons labeled happiness, disgust, sadness, fear, anger, surprise, or neutral located at the bottom of the photos.

For the training part, the participants had to perform a categorization task on six images of emotional facial expressions and one neutral image. The emotional images were presented with an expressive intensity of 100%.

The experimental part had the same characteristics as the training part. Pictures of the training section were not reused. Each of the 42 images, described in the materials section, were presented at 100% level of intensity as configured in the software in accordance with the FACS definition (Ekman et al., 2002). Images were exposed in random order. The images of

the modeled faces were displayed until the category was selected by the participant (happiness, disgust, sadness, fear, anger, surprise, and neutral). The choice of category led to the continuation of the experiment.

For each participant, a mean percent correct score was calculated. There was no feedback during the training and experimental part to avoid the learning reinforcement process. At the end of the experiment, demographic information was collected. This concerned gender, age, socio-professional category, and professional activity.

Results

The total categorization score for all participants was 85.5%. As some papers suggest gender differences in judgement of emotional facial expressions (Birditt & Fingerman, 2003; Hall & Matsumoto, 2004; Ryan & Gauthier, 2016), a Welch two-sample t test was applied to gender and the total score of each participant. There was no significant difference ($p = .73$, $d = 0.11$, power level = .99) between men ($M = 84.85$, $SD = 8.47$) and women ($M = 85.83$, $SD = 8.85$).

A three-way analysis of variance was conducted with gender as a between participant factor, and face genders of stimuli (man, woman) and expressions (happiness, disgust, sadness, fear, anger, surprise, neutral) as repeated factors (see Fig. 8). There was no effect involving participants' gender. A Tukey HSD test was used to examine significant effects of the gender and emotional category of stimuli. There was a main effect of
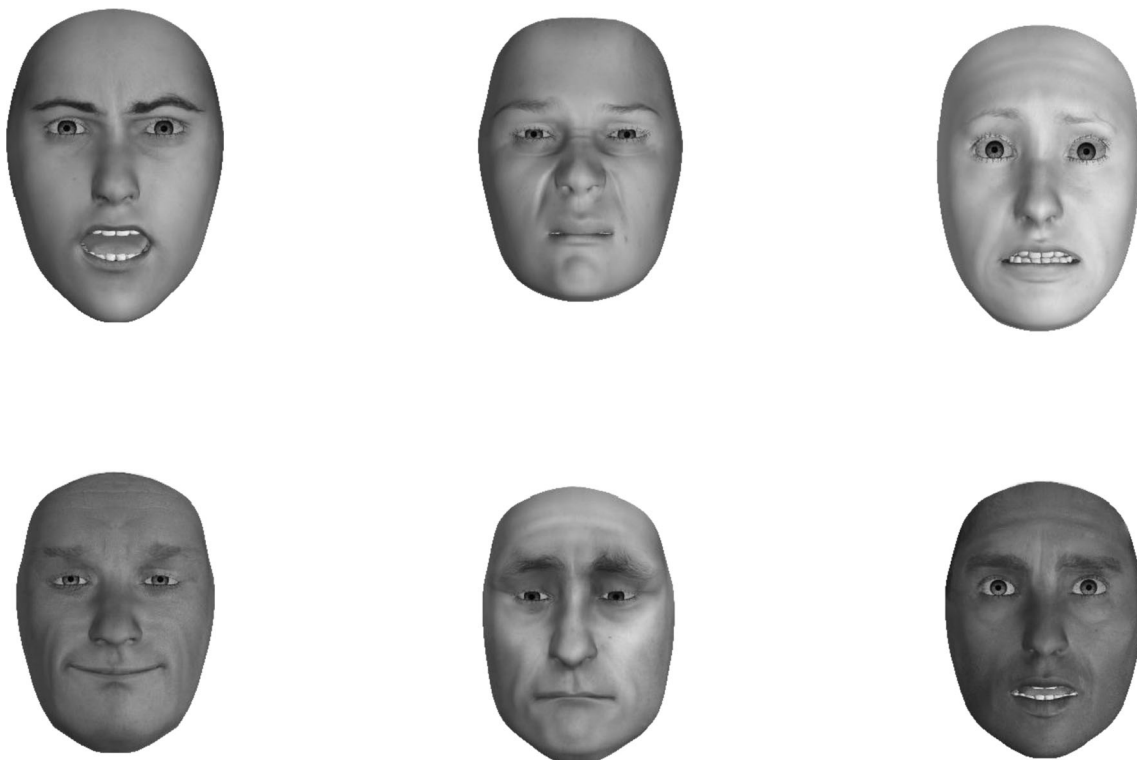

Fig. 8 Models used for the categorization task (ex: anger, disgust, fear, happiness, sad, surprise)

face gender, $F(1, 6) = 10.76$, $MSq = 9.32$, $p < .001$, $\chi^2 = .02$. Women's faces ($M = 88.53\%$, $SE = 1.32$) were better categorized than men's faces ($M = 82.47\%$, $SE = 1.64$). There was a main effect with emotional configurations, $F(1, 6) = 9.49$, $MSq = 8.22$, $p < .001$, $\chi^2 = .08$. Anger ($M = 96.21\%$, $SE = 1.17$) and neutral ($M = 93.94\%$, $SE = 1.48$) facial configurations were categorized more accurately than others. Fear was the least categorized expression ($M = 73.86\%$, $SE = 2.71$; all $ps < .01$). Happiness ($M = 79.92\%$, $SE = 2.46$), disgust ($M = 86.74\%$, $SE = 2.09$), sadness ($M = 85.61\%$, $SE = 2.15$) and surprise ($M = 82.2\%$, $SE = 2.34$) were well categorized.

The type of errors made by participants were examined (Table 3).

Anger (96.21%) was the most recognized emotional configuration with less than 2% confused with fear and surprise. Happiness (79.92%) was the most confused with neutral (18.18%), disgust (86.74%) with anger (10.61%), and sadness (85.61%) with neutral (9.85%). Surprise (82.20%) and fear (73.86%) were confused the most (16.67%), and neutral (93.94%) was confused with sadness (3.41%).

To compare our results (Table 3) with the POFA database (Table 4), we used the confusion matrix table provided in the Cigna et al. (2015) paper on the POFA's pictures of emotional facial expressions. These tables describe the categorization performance and the error rates of participants for each emotional category. The diagonal results correspond to the percentage of correct responses to the facial expressions presented. For each emotional facial expression, rows indicate the percentage of confusion with others' emotional facial expressions. As results for neutral expression were not present in this table, neutral FACSHuman expression scores were not used in statistical computation. A Welch unpaired two-sample t test was used to compare the overall recognition score between POFA and FACSHuman for the six basic facial expressions. The results indicated that there was no significant difference ($p = .44$, $d = 0.47$, $power = .65$) between POFA ($M = 78.93$, $SE = 5.51$) and FACSHuman ($M = 84.08$, $SE = 3.07$) total score.

Our data were significantly different from normal distribution for each category. One-sample Wilcoxon rank-sum tests were used to compare the results obtained for the

FACSHuman expressions with those from the POFA, in a one-by-one categorical comparison. The results showed that FACSHuman expression scores were significantly higher except for happiness and surprise expressions, where the POFA expressions were better rated (see Fig. 9 and Table 5). Effects size showed large differences.

Comparison with previous works. As we have seen in the introduction, earlier software programs which are the most similar to FACSHuman are HapFACS and FACSgen. We computed means and standard deviations with the data found in the HapFACS studies for static emotional expressions at their respective maximum intensity. The overall recognition rate for the six basic emotional facial expressions created with FACSHuman was 84.09 ($SD = 7.51$). This result is situated between the results of the two previously cited software programs for the same emotions (Table 6).

Discussion

The aim of this study was to evaluate images generated by FACSHuman to reproduce facial expressions of emotion as defined by theoretical models (Ekman, 1971, 1992; Ekman & Oster, 1979). The results obtained were comparable to those of the various picture databases such as POFA (Cigna et al., 2015; Ekman, 1976; Palermo & Coltheart, 2004). Analyses of categorization data collected showed that modeled face creation for experimental use was possible with the FACSHuman plugins. Overall, the categorization scores of FACSHuman emotional facial expressions rendered with women's morphological criteria were better than those for men's faces. Anger and neutral facial configurations were the most accurately categorized (Ekman et al., 2002). Fear was the least categorized expression and confused with surprise (Becker, 2017). This result is probably due to the AUs shared between these two expressions (Du & Martinez, 2015).

There were no significant differences between the results of the POFA evaluation (Cigna et al., 2015) and the FACSHuman generated pictures. The overall percentage of categorization rate was high and related to those found in other database validation studies (Goeleven et al., 2008; Langner

Table 3 Confusion matrix of EFE (in %) for FACSHuman pictures

	Happiness	Disgust	Sadness	Fear	Anger	Surprise	Neutral
Happiness	79.92	1.52	.00	0.00	0.00	0.38	18.18
Disgust	0.38	86.74	1.52	0.00	10.61	0.00	0.76
Sadness	0.00	2.65	85.61	0.76	0.38	0.76	9.85
Fear	0.00	0.76	0.38	73.86	7.95	16.67	0.38
Anger	0.00	0.38	0.00	1.89	96.21	1.52	0.00
Surprise	0.00	0.00	0.76	14.39	2.27	82.2	0.38
Neutral	0.00	0.00	3.41	0.00	1.89	0.76	93.94

Table 4 Confusion matrix of EFE (in %) for POFA pictures from (Cigna et al., 2015)

	Happiness	Disgust	Sadness	Fear	Anger	Surprise
Happiness	97.4	0.4	1.0	0.3	0.1	0.8
Disgust	0.4	68.4	0.4	0.1	29.9	1.0
Sadness	0.4	12.1	72.9	7.4	3.5	3.7
Fear	0.1	4.1	0.7	62.3	0.6	32.2
Anger	0.4	4.7	1.8	3.6	82	7.6
Surprise	0.9	0.8	0.2	7.1	0.5	90.6

Table 5 Categorical scores comparison between POFA and FACSHuman

	Happiness	Disgust	Sadness	Fear	Anger	Surprise
POFA	97.4	68.4	72.9	62.3	82	90.6
FACSHuman	79.92	86.74	85.61	73.86	96.21	82.2
p	< .001	< .001	< .001	< .001	< .001	< .001
r	.88	.90	.89	.88	.94	.89

et al., 2010). The computer modeling of experimental research material, via the use of FACSHuman, offers possibilities in relation to the creation of a photographic database of directed actors for emotional facial expressions.

The material was produced with the FACSHuman software program and compared with the one found and described in the FACS manual (Ekman et al., 2002).

Experiment 2

The FACS describes facial expression by the use of codification of Facial Action Units (AUs). The FACS manual is a tool to train people to recognize facial movements that involve one or more AUs. In this experiment, an evaluation of the accuracy of unique facial movements identified as AUs was conducted by non-FACS coders in order to compare them with those described and found in the FACS manual.

Method

Participants Twenty-two women (M_{age} = 39.09, SD_{age} = 11.71) and 28 men (M_{age} = 47.37, SD_{age} = 11.37) were engaged in this experiment via the LinkedIn professional network. They were recruited by the internal messaging system. The participants validated, via a checkbox, a form giving free, informed, and express consent before the experiment. The experiment could not begin without validation of the form. It took place on an Internet browser on the participant's computer and respected people's anonymity.

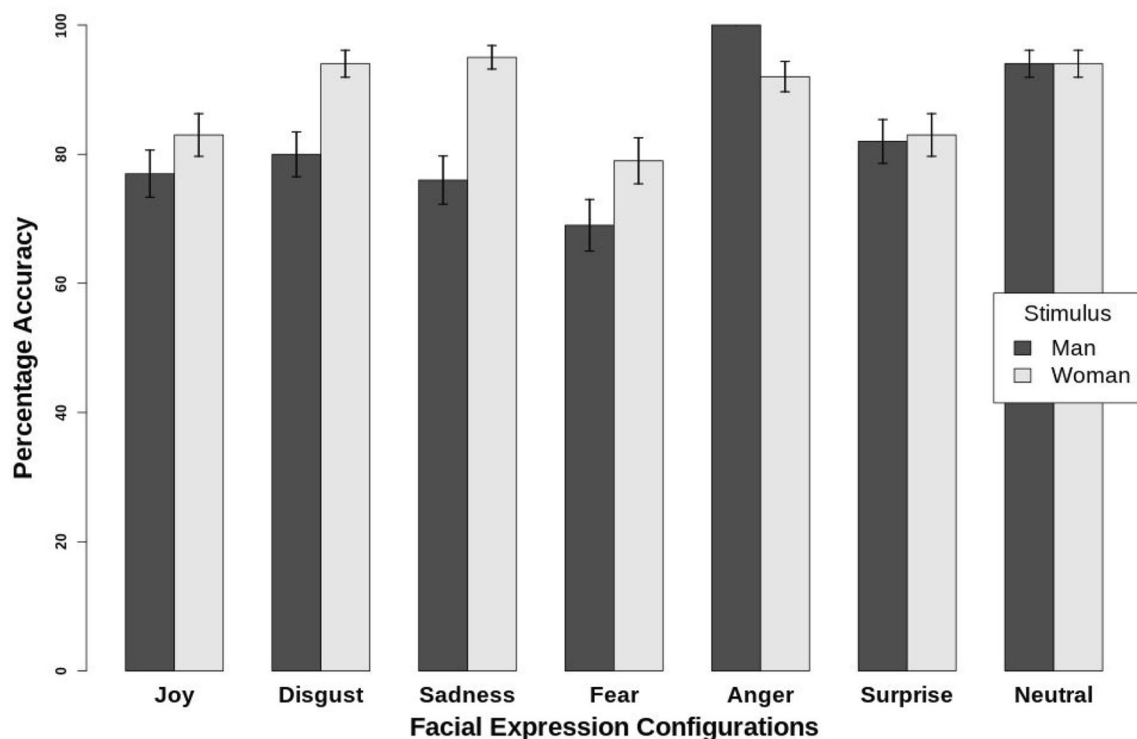**Materials** The experimental material consisted of 26 photos from the FACS reference manual (Ekman et al., 2002) and

Fig. 9 Percentage of categorization by stimulus gender for FACSHuman Stimuli. *Error bars* represent standard errors

Table 6 Comparison of the recognition rate by software

Emotion	HapFACS	FACSgen	FACSHuman
Anger	75.0	87.82	96.21
Disgust	75.0	68.59	86.74
Fear	87.5	72.44	73.86
Happiness	92.5	88.46	79.92
Sadness	90.0	83.97	85.61
Surprise	97.5	87.82	82.2
Means (*SD*)	86.25 (9.32)	81.52 (8.75)	84.09 (7.51)

Note. Data from HapFACS studies for HapFACS & FACSgen columns (Amini et al., 2015) Appendix Table 14

26 pictures generated with FACSHuman software program. These photos represented the same Caucasian man from the FACS and the same modeled Caucasian man for the FACSHuman software program.

For both models, the 25 pictures represent one AU at once, plus a photograph of the face with a neutral expression. These last were added to test faces with no AUs in action, and to increase the validity of participants' responses (Russell, 1993), one from the FACS manual, while the other was generated with the FACSHuman software program. A total of 52 trials were presented to the participants. Twenty-six were congruent, where both models expressed the same AUs, and twenty-six were non-congruent, where the two faces expressed different AUs. Appendix A4 presents the 26 AUs alone for the whole picture presented.

The following experiment was constructed using the JSPsych Experiment Creation Framework (de Leeuw, 2015).

Procedure First, a brief presentation of the theme of the experiment and the researcher was made. Then came the presentation of the experiment and its average duration. Following this presentation, participants were instructed to sign the informed consent form. All participants were volunteers and were free to withdraw at any stage of the study. No remuneration was given. Written informed consent was signed by the participants.

The training session and main experiment consisted in a comparison task of two facial expression images displayed side by side, one from the FACS manual and one created with FACSHuman software. Participants had to complete the sentence "The expressions of these two people are" by choosing the word "Different" or "Identical". Images exposed only one Action Unit at a time. They were presented with an expressive intensity of 100% according to the FACS manual references. The coupled images were exposed in random order. They were displayed until the participant selected one word. The choice of a word led to the continuation of the experiment. There was no time limitation for completing the experiment.

The training sessions consisted of one congruent and one incongruent trial. For the experimental parts, 26 congruent and 26 non-congruent trials were displayed in random order. Each pair was seen only once.

During the training and the experiment part, transition from one instruction screen to another was effected by pressing a "next" button displayed on the screen. At the end of the experiment, demographic information was collected. This concerned gender and age. The data was processed statistically using a signal discrimination procedure. The results were subjected to an analysis with regard to the performance measurement procedures for same-different experimental protocol as described by Macmillan and Creelman (2005).

Results

For all stimuli, participants had a discrimination rate of 82.44% ($SD = 0.09$) (Appendix Table 14). As for the previous experiment, we tested gender differences. There was no difference between men ($M = 82.41$, $SD = 9.02$) and women ($M = 82.49$, $SD = 8.83$) in the results obtained for the task (Welch two-sample t test, $p = 0.974$, $d = 0.01$, power level = .99).

The least recognized AUs were *lip presser* (AU 24, 38%) and *nasolabial furrow deepener* (AU 11, 46%). The most recognized AUs were *Eye closure* (AU 43, 98%) and *Upper lid raiser* (AU 5, 96%) (Table 7).

The results showed a good discrimination of the stimuli presented to the participants ($A' = 0.90$) as well as a good sensitivity ($d' = 1.97$). However, participants presented a light conservative response bias ($C = 0.30$). Participants were more inclined to respond "different" to the question asked.

Discussion

In this experiment, the accuracy of the most used AUs was evaluated by non-FACS coders. They were engaged in a same-different task to examine the accuracy of FACSHuman generated AUs expressed one at a time compared with those found in the FACS manual.

Results showed a good discrimination rate compared to previous studies (Sayette et al., 2001). Our results showed that an accuracy of 82.44% was consistent with previous findings that reported an average recall of 81% for Amini et al. (2015) and 90% interraters' reliability for Krumhuber et al. (2012) on single AUs by FACS coders expressed at their highest intensity levels. The low discrimination rate of AU 24 (lip presser) and AU 11 (Nasolabial Furrow Deepener) could be explained by the nature of the stimuli, which were pictures rather than videos. However, as pointed out in the FACS manual, also by the small facial movements these two AUs involved the face.

These two AUs are very small movements of the face contrary to AU 43 (Eye closure) and AU 5 (Upper lid raiser), which are movements of great amplitude. The results obtained

Table 7 Percentage of good response by action units for congruent and non-congruent trials

AU FH		Congruent	Non-congruent
1	Inner Brow Raise	76	86
2	Outer Brow Raise	78	96
4	Brow Lowerer	80	98
5	Upper Lid Raise	96	94
6	Cheek Raise	68	98
7	Lids Tight	80	74
9	Nose Wrinkle	68	84
10	Upper Lip Raiser	80	86
11	Nasolabial Furrow Deepener	46	100
12	Lip Corner Puller	92	98
13	Sharp Lip Puller	90	96
14	Dimpler	78	86
15	Lip Corner Depressor	86	96
17	Chin Raiser	92	96
18	Lip Pucker	84	82
20	Lip Stretch	86	66
23	Lip Tightener	72	100
24	Lip Presser	38	94
25	Lips Part	70	96
26	Jaw Drop	90	94
27	Mouth Stretch	66	88
28	Lip Suck	92	94
29	Jaw Thrust	64	90
35	Cheek Suck	88	84
43	Eye Closure	98	84

in the first and second experiment by non-professional FACS coders were good. They provided a good evaluation of the accuracy of the experimental material produced with the FACSHuman software program in comparison to the FACS.

In the next two experiments, two certified FACS coders were recruited to carry out an evaluation of the accuracy of AUs alone and AU combinations. The two previous studies were conducted to evaluate the likelihood of combinations and AUs alone available in the software by non-coders. These two additional experiments were conducted to provide a technical evaluation of the accuracy of AU movements produced by the software in comparison to the FACS.

Experiment 3

In the previous experiments, evaluation of emotional facial expressions (Experiment 1) and AUs alone (Experiment 2) was conducted by non-FACS coders. In order to be used as experimental material, images and videos produced by

FACSHuman need to be as compliant as possible with the FACS. This compliance will facilitate the communication and the description of the facial configurations used as research stimuli.

In this experiment, we used a similar protocol described in the Amini et al., 2015 article and the materials produced with the FACSHuman software program submitted to certified FACS coders. However, we did not extend the protocol by the use of different intensities; we only used 100% intensity as described in the FACS manual. This consisted of the encoding by FACS coders of a series of pictures and videos generated with FACSHuman showing only one AU at a time. AU combinations were evaluated in the next experiment.

Method

Participants Two men participated in this study. They were recruited via an advertisement on the LinkedIn professional network.

The first coder was 39 years old and had 4 years' experience in FACS coding. The second coder was 30 years old and had 5 years' experience in FACS coding. The second coder had had regular facial coding practice using the FACS methodology as part of his doctoral research in nonverbal communication. Participants validated a consent statement prior to the start of the experiment.

Materials The experimental material consisted of 47 images and 47 videos (see Appendix A5) of the action units described in the FACS manual (Ekman et al., 2002) expressed by a Caucasian man. The tested AUs were the 26 AUs used in the previous experiment with the addition of 21 units. These last AUs were the miscellaneous actions and supplementary codes. These are extended AUs or movement descriptors described in the FACS manual (Ekman et al., 2002) and mostly involved in the FACS coding procedure.

Procedure For each participant, images of facial expressions that represent the AUs presented with an expressive intensity of 100% were displayed on the screen one at a time. Participants could freely select the different images to encode by selecting them with tabs at the bottom of the screen, which were randomly distributed.

For each image, an image of the face with a neutral expression was presented side by side on the right of the screen with the image that the coder had to analyze. A video of the action unit was also available for the coder to the right of the neutral one. Videos were playable as many times as needed. Participants were asked to code the different images and videos presented on the screen according to the notation system defined in the FACS (Ekman et al., 2002). They entered their FACS code in a text box below the images. There was no time limitation and participants could perform the coding task freely in multiple sessions as described in the FACS coding procedure.

Results

The results of the encoding procedure for the AUs were analyzed using the α-Agreement coding procedure of Krippendorff (2004, 2011). The agreement coefficient computed was moderate (α = 0.59). The first coder obtained a recognition score of 63.83% and the second coder a score of 85.11% (Appendix Table 15).

Discussion

This experiment was conducted to evaluate the compliance of AUs, displayed by a model created with FACSHuman, with the FACS. This evaluation of images and videos was performed by an encoding task conducted by two certified FACS coders. The weaker result produced by the first coder, as compared with the second one, can be explained by their level of regular practice, which was lower than that of the second coder. AU 6 (Cheek Raise), 13 (Sharp Lip Puller), and 25 (Lips Part) were not coded correctly by either coder. For AU 6 this detail was discussed with the coders. It was due to the lack of precision of the wrinkles at the corners of the eyes, which is a criterion for coding this AU. This detail was corrected in the software after data analysis of Experiments 3 and 4. AU 13 was coded as 12, which is the most similar movement. Number 25 is involved in different facial configurations and can be coded differently in accordance with its intensity. The FACS certification requires natural and spontaneous facial expressions to be coded into movements. In the next experiment, the two FACS coders performed a codification task of complex facial expressions with multiple AUs generated with the FACSHuman software program.

Experiment 4

Complex facial expressions are composed of the activity of multiple AUs at one time. In this experiment, certified FACS coders were asked to code 54 facial configurations described in the FACS manual. The aim of this experiment was to evaluate the accuracy of complex facial expression images produced with the FACSHuman software program.

Method

Participants Participants were the same as in the previous experiment.

Materials Experimental material consisted of 54 combinations of AUs (Appendix Table 13). The interface was the same as the one described in the previous experiment. The computer screen showed the face of a Caucasian man. Participants had at their disposal a neutral image of the face as well as an image of the combination and the corresponding video corresponding. It displayed in succession the coding instructions and the different combinations of AUs to be coded in a randomized and anonymous sequence. Coders always had a reminder sheet containing all the codes of the FACS repository (Ekman et al., 2002) as described in the FACS encoding procedure.

Procedure In this experiment, the interface and protocol were the same as the one described in the previous experiment. The participants were asked to code each expression presented on the screen with an expressive intensity of 100%. The results were reported below the images using the codification system.

Statistical analyses The scoring procedure for calculation of the agreement index was the one described and used in Ekman et al. (1972) from Wexler (1972) to analyze the results obtained in the coding sessions.

This index is computed with the following formula:

Index of agreement

$$= \frac{(number\ of\ AUs\ on\ which\ coder\ 1\ and\ 2\ agreed)\ x\ 2}{total\ number\ of\ AUs\ scored\ by\ the\ two\ coders}$$

Results and discussion

In this experiment, we tested the accuracy of AU combinations with two certified FACS coders. A result of .68 constitutes a satisfactory intercoder agreement (Ekman et al., 1972). The recognition averages for the AUs were 68.08% for the first coder and 81.56% for the second coder (Appendix Table 16). As for the previous experiment, there was a difference in recognition between the two coders.

General discussion

This article describes and presents the evaluation of the FACSHuman software program. To carry out this evaluation, we conducted four experiments to assess the accuracy and the compliance of facial expressions and AUs with the FACS. In the first and second experiments, lay persons had to evaluate emotional expressions as well as unique AUs produced with FACSHuman. The first experiment was a categorization task of six basic emotional facial expressions. The second was a same-different comparison task of AUs expressed one at a time between the one found in the FACS manual and the one produced by FACSHuman. In the third and fourth experiments, we investigated the recognition of single AUs and AU combinations with certified FACS coders. FACS coders were trained to identify combinations and unique AUs as well as some EmFACS emotional configurations. However, lay people do not have such skills. Thus, we decided to present lay people with AUs alone for some participants, and with

emotional facial expressions for other participants. These two groups of participants allowed us to increase the evaluation panel and the ecological validity.

The first experiment reported an 85.5% recognition rate by participants for emotional expressions. The second experiment showed a good accuracy of AUs only presented to the lay participants and related to the one found in the FACS manual. In the third and fourth experiments, 47 AUs and 54 AU combinations were evaluated by two certified FACS coders. Results showed moderate-to-satisfactory intercoder agreement.

Nevertheless, our results show that this software does not dispense with performing an evaluation process for the sets of images or videos produced. Indeed, as for the image banks previously produced by different research laboratories, the perception of emotional facial expressions presents a minimal variance that cannot be predicted by the aesthetic and morphological criteria chosen during the creation of the avatars. However, FACSHuman frees the researcher from the selection and the direction of actors. Evaluation can now be done quickly via research application environments made available to participants on the Internet. With this way of evaluation, the experimenter is free to choose groups and segments by age, gender, and social groups ... to which they wish to submit this evaluation. This ensures greater calibration for the image set produced for the population targeted by the experiment. The results obtained, in a categorization task, showed that the images from the available data banks offer a recognition rate inferior or comparable to that of digitally created images.

Limits of the present work

In the first experiment, we did not use animation as compared to some of the other validation protocols. Further validations, such as levels of believability, realism, or intensity, may be required to evaluate facial animations. Another topic could be the production and evaluation of micro-expressions for which users could control the speed and intensity of the action units used in the created facial expression. A comparative study, conducted by Krumhuber et al. (2017), described the key dimensions and properties of dynamic facial expression datasets available for researchers. The authors highlighted the need for an evaluative study of the material used such as categorization or judgement tasks and the accuracy of the conveyed emotions. It also pointed out the use of FACS as a tool "of considerable value", as it can be used to compare observed facial expressions with reported emotions or physiological responses. Indeed, the different morphological configurations, skin tints, and spatial configurations of the elements characterizing a human face are likely to induce emotional interpretation of facial expression (Palermo & Coltheart, 2004). Reciprocally, emotions can affect gender categorization (Roesch et al., 2010) as well as social appraisal (Mumenthaler & Sander, 2012). This effect is found in the

results obtained in the first two studies. The different avatars were not evaluated in the same way by the participants although they met the same characteristics of neutrality. These elements must then be considered as variables of significant confusion when creating an experiment. The equipment must be calibrated before being used in an experiment. The results obtained can be used as comparators (expected results).

Declaration

Competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Tables

The order of AUs presented in the tables below follow the presentation order of the Facial Action Coding System manual.

Table 8   Usable Action Units in FACSHuman

Action units	Simple	Right / Left	Alternative
Upper Face AUs			
1 Inner Brow Raise	X	X	X
2 Outer Brow Raise	X	X	X
4 Brow Lowerer	X	X	X (41,42,44)
5 Upper Lid Raise	X	X	
6 Cheek Raise	X	X	X (2)
7 Lids Tight	X	X	
43 Eye Closure	X	X	
45 Blink	Use 43	Use 43	
46 Wink	Use 43	Use 43	
70 Brows Not Visible	-	-	-
71 Eyes Not Visible	-	-	-
Head Positions			
51 Turn Left	X		
52 Turn Right	X		
53 Head Up	X		
54 Head Down	X		
55 Tilt Left	X		
56 Tilt Right	X		
57 Forward	X		
58 Back	X		
Eye Positions			
61 Eyes Left	X	X	
62 Eyes Right	X	X	
63 Eyes Up	X	X	
64 Eyes Down	X	X	
65 Walleye	X		
66 Cross-eye			
Lip Parting and Jaw Opening			
25 Lips Part	X	X	X
26 Jaw Drop	X		X(4)
27 Mouth Stretch	Use 26		

Note. (X) usable in the plugin; (-) non implemented

Table 9   Usable action units in FACSHuman

Action units	Simple	Right / Left	Alternative
Lower Face AUs			
9 Nose Wrinkle	X	X	X
10 Upper Lip Raiser	X + 25	X	
11 Nasolabial Furrow Deepener	X	X	
12 Lip Corner Puller	X	X	X(2)
13 Sharp Lip Puller	X	X	
14 Dimpler	X	X	
15 Lip Corner Depressor	X	X	
16 Lower Lip Depress	X	X	
17 Chin Raiser	X		X
18 Lip Pucker	X		X
20 Lip Stretch	X		
22 Lip Funneler	X		X(2)
23 Lip Tightener	X		
24 Lip Presser	X		
28 Lips Suck	X		X(2)
72 Lower Face Not Visible	-	-	-
Miscellaneous AUs			
8 Lips Toward Each Other	-	-	-
19 Tongue Show	X + 26		
21 Neck Tightener	-	-	-
29 Jaw Thrust	X		
30 Jaw Sideways	X		
31 Jaw Clencher	X		
32 Bite	X		
33 Blow	X		
34 Puff	X		
35 Cheek Suck	X		
36 Tongue Bulge	-	-	-
37 Lip Wipe	-	-	-
38 Nostril Dilate	X		
39 Nostril Compress	X		

Note. (X) usable in the plugin; (-) non implemented; (+) usable with alternative Action Units (number)

Table 10   Additional action units usable in FACSHuman

FACSHuman	Simple	Right / Left	Alternative
pupil dilatation	X	X	
pupil contraction	X	X	

Note. (X) usable in the plugin; (-) non implemented

Table 11 Action units used in Experiment 2

Action unit number	Action unit / descriptor
-	Neutral face
1	Inner brow raiser
2	Outer brow raiser
4	Brow lowerer
5	Upper lid raiser
6	Cheek raise
7	Lid tightener
9	Nose wrinkler
10	Upper lip raiser
11	Nasolabial furrow deepener
12	Lip corner puller
13	Sharp lip puller
14	Dimpler
15	Lip corner depressor
17	Chin raiser
18	Lip pucker
20	Lip stretcher
23	Lip tightener
24	Lip presser
25	Lips part
26	Jaw drop
27	Mouth stretch
28	Lip suck
29	Jaw thrust
35	Cheek tuck
43	Eye closure

Table 12 Action units used in Experiment 3

Action unit number	Action unit / descriptor
	Neutral face
1	Inner brow raiser
2	Outer brow raiser
4	Brow lowerer
5	Upper lid raiser
6	Cheek raiser
7	Lid tightener
9	Nose wrinkler
10	Upper lip raiser
11	Nasolabial deepener
12	Lip corner puller
13	Sharp lip puller
14	Dimpler
15	Lip corner depressor
16	Lower lip depressor
17	Chin raiser
18	Lip pucker
20	Lip stretcher
22	Lip funneler
23	Lip tightener
24	Lip pressor
25	Lips part
26	Jaw drop
29	Jaw thrust
31	Jaw clencher
32	[Lip] bite
33	[Cheek] blow
34	[Cheek] puff
35	[Cheek] suck
38	Nostril dilator
39	Nostril compressor
43	Eyes closed
51	Head turn left
52	Head turn right
53	Head up
54	Head down
55	Head tilt left
56	Head tilt right
57	Head forward
58	Head back
61	Eyes turn left
62	Eyes turn right
63	Eyes up
64	Eyes down
65	Walleye
66	Cross-eye

Table 13 Action unit combinations used in Experiment 4

Number	FACS combination codes
1	1 + 2
2	1 + 4
3	1 + 2 + 4
4	1 + 2 + 5
5	4 + 5
6	5 + 7
7	6 + 43
8	6 + 7 + 12
9	6 + 12 + 15
10	6 + 12 + 15 + 17
11	6 + 12 + 17 + 23
12	7 + 12
13	7 + 43
14	9 + 17
15	9 + 16 + 25
16	10 + 14
17	10 + 15
18	10 + 12 + 25
19	10 + 17
20	10 + 15 + 17
21	10 + 16 + 25
22	10 + 17 + 23
23	10 + 20 + 25
24	10 + 23 + 25
25	10 + 12 + 16 + 25
26	12 + 15
27	12 + 17
28	12 + 23
29	12 + 24
30	12 + 25 + 26
31	12 + 25 + 27
32	12 + 15 + 17
33	12 + 16 + 25
34	12 + 17 + 23
35	20 + 23 + 25
36	22 + 23 + 25
37	23 + 25 + 26
38	14 + 17
39	14 + 23
40	15 + 17
41	15 + 23
42	17 + 23
43	17 + 24
44	18 + 23
45	20 + 25 + 26
46	20 + 25 + 27
47	4 + 5 + 7 + 24
48	10 + 16 + 25 + 26
49	14 + 54 + 62 + 64
50	1 + 2 + 4 + 5 + 20 + 25 + 26
51	12 + 53 + 64
52	6 + 12
53	1 + 4 + 15
54	1 + 2 + 5 + 25 + 27

Table 14 Signal detection values by participants for Experiment 2

Participant	Good Reponses	Hit Rate	False Alarm Rate	*d'*	beta	*a'*	*bdd'*
1	.85	.86	.15	2.09	0.95	.91	-.04
2	.89	.89	.12	2.44	0.95	.94	-.04
3	.85	.75	.04	2.44	3.81	.92	.79
4	.87	.79	.04	2.56	3.49	.93	.74
5	.87	.75	.02	2.74	6.79	.93	.89
6	.8	.82	.23	1.66	0.86	.87	-.16
7	.83	.71	.04	2.33	4.07	.91	.82
8	.74	.64	.15	1.39	1.57	.83	.51
9	.76	.79	.27	1.41	0.88	.84	-.15
10	.85	.75	.04	2.44	3.81	.92	.79
11	.93	.86	.02	3.14	4.82	.96	.79
12	.89	.79	.02	2.86	6.23	.94	.87
13	.74	.64	.15	1.39	1.57	.83	.51
14	.85	.79	.08	2.22	2.02	.92	.53
15	.87	.96	.23	2.54	0.26	.93	-.78
16	.80	.68	.08	1.89	2.48	.88	.70
17	.85	.75	.04	2.44	3.81	.92	.79
18	.76	.61	.08	1.70	2.66	.86	.77
19	.89	.86	.08	2.49	1.56	.94	.33
20	.96	.96	.04	3.57	0.94	.98	-.04
21	.78	.68	.12	1.66	1.84	.87	.57
22	.80	.71	.12	1.76	1.75	.88	.51
23	.83	.82	.15	1.94	1.10	.90	.09
24	.72	.82	.38	1.21	0.68	.81	-.48
25	.91	.82	.02	2.99	5.57	.95	.83
26	.74	.50	.02	2.07	8.52	.86	.96
27	.85	.93	.23	2.20	0.45	.91	-.59
28	.63	.68	.42	0.66	0.92	.70	-.22
29	.81	.75	.12	1.87	1.63	.89	.44
30	.59	.21	.02	1.28	6.23	.78	.99
31	.78	.79	.23	1.53	0.96	.86	-.05
32	.74	.64	.15	1.39	1.57	.83	.51
33	.85	.79	.08	2.22	2.02	.92	.53
34	.78	.82	.27	1.54	0.79	.86	-.26
35	.94	.96	.08	3.23	0.54	.97	-.38
36	.85	.86	.15	2.09	0.95	.91	-.04
37	.72	.50	.04	1.77	4.78	.85	.92
38	.89	.79	.02	2.86	6.23	.94	.87
39	.81	.68	.04	2.23	4.29	.90	.84
40	.89	.82	.04	2.69	3.13	.94	.69
41	.61	.36	.12	0.83	1.92	.74	.86
42	.91	.82	.02	2.99	5.57	.95	.83
43	.81	.79	.15	1.81	1.23	.89	.20
44	.93	.86	.02	3.14	4.82	.96	.79
45	.65	.36	.04	1.40	4.47	.81	.96
46	.89	.79	.02	2.86	6.23	.94	.87
47	.94	.96	.08	3.23	.54	.97	-.38
48	.89	.86	.08	2.49	1.56	.94	.33
49	.98	.96	.02	3.87	1.68	.99	.31
50	.83	.75	.08	2.10	2.20	.91	.60

Table 15 Encoded response by the two coders for Experiment 3

FACSHuman Action Units	Coder 1	Coder 2
1	1	1
2	2	2
4	4	4
5	5	5
6	10	13
7	7	7
9	9	9
10	10	10
11	10	11
12	12	12
13	12	12
14	12	14
15	15	15
16	16	26
17	17	17
18	22	18
20	23	20
22	25	22
23	24	23
24	24	24
25	**10**	**10**
26	26	27
28	28	24
29	26	29
31	23	31
32	32	32
33	34	33
34	33	34
35	35	35
38	38	38
39	39	39
43	45	43
51	51	51
52	52	52
53	58	53
54	57	54
55	55	55
56	56	56
57	57	57
58	57	58
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66

Table 16 Encoded response by the two coders for Experiment 4

FACSHuman Combination	Coder 1	Coder 2
1+2	1+2	1+2
1+4	1+4	1+4
1+2+4	1+4	1+2+4
1+2+5	1+2+5	1+2+5
4+5	4+5	4+5
5+7	5+7	5+7
6+43	10+43	6+43
6+7+12	6+12+23	6+12
6+12+15	6+10+12	6+12+17
6+12+15+17	6+10+12	6+12+17+24
6+12+17+23	6+12+17	6+17
7+12	7+12	7+12
7+43	43	7+43
9+17	9+10+15	9+17
9+16+25	9+10+25	9+25
10+14	11+12+14	10+14
10+15	10+15	10+15
10+12+25	11+12+25	10+12
10+17	11+17	10+17
10+15+17	10+15	10+15+17
10+16+25	10+16+25	10+25
10+17+23	10+17+24	10+17
10+20+25	10+23+25	10+20+25
10+23+25	10+25	10+23+25
10+12+16+25	10+12+25	10+12+25
12+15	12+23	17+20
12+17	12+17	12+17
12+23	12+17+24	12+23
12+24	12+17+24	12+24
12+25+26	12+25+26	12+26
12+25+27	25+26	12+27
12+15+17	17+23	17+20
12+16+25	12+25	12+25
12+17+23	12+24	12+17
20+23+25	23+25	20+23+25
22+23+25	18+24+25	22+23
23+25+26	25+26	23+26
14+17	12+24	17+24
14+23	12+24	14+23
15+17	15+17	15+17
15+23	15+24	15+23
17+23	17+24	17+23
17+24	17+24	17+24
18+23	18+24	18+23
20+25+26	23+25+26	20+26
20+25+27	23+25+26	20+27
4+5+7+24	4+5+7+24	4+5+7+23
10+16+25+26	10+25+26	10+26
14+54+62+64	12+54+62	54+62
1+2+4+5+20+25+26	1+2+5+23+25+26	1+2+4+5+20+26
12+53+64	12+53+64	12+53
6+12	6+12	6+12
1+4+15	4+15	1+4+15
1+2+5+25+27	1+2+5+25+26	1+2+5+26

References

Amini, R., Lisetti, C., & Ruiz, G. (2015). HapFACS 3.0: FACS-Based Facial Expression Generator for 3D Speaking Virtual Characters. *IEEE Transactions on Affective Computing*, 6(4), 348–360. https://doi.org/10.1109/TAFFC.2015.2432794

Aneja, D., Colburn, A., Faigin, G., Shapiro, L., & Mones, B. (2017). Modeling Stylized Character Expressions via Deep Learning. In: S.-H. Lai, V. Lepetit, K. Nishino, & Y. Sato (Eds.), *Computer Vision – ACCV 2016* (Vol. 10112, pp. 136–153). https://doi.org/10.1007/978-3-319-54184-6-9

Bänziger, T., Grandjean, D., & Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: The Multimodal Emotion Recognition Test (MERT). *Emotion*, 9(5), 691–704. https://doi.org/10.1037/a0017088

Baudouin, J.-Y., & Humphreys, G. W. (2006). Configural Information in Gender Categorisation. *Perception*, 35(4), 531–540. https://doi.org/10.1068/p3403

Becker, D. V. (2017). Facial gender interferes with decisions about facial expressions of anger and happiness. *Journal of Experimental Psychology: General*, 146(4), 457–463.

Bennett, C. C., & Šabanović, S. (2014). Deriving Minimal Features for Human-Like Facial Expressions in Robotic Faces. *International Journal of Social Robotics*, 6(3), 367–381. https://doi.org/10.1007/s12369-014-0237-z

Birditt, K. S., & Fingerman, K. L. (2003). Age and Gender Differences in Adults' Descriptions of Emotional Reactions to Interpersonal Problems. *The Journals of Gerontology: Series B*, 58(4), P237-P245. https://doi.org/10.1093/geronb/58.4.P237

Bruce, V., Burton, A. M., Hanna, E., Healey, P., Mason, O., Coombes, A., Fright, R., & Linney, A. (1993). Sex Discrimination : How Do We Tell the Difference between Male and Female Faces? *Perception*, 22(2), 131-152. https://doi.org/10.1068/p220131

Bruyer, R., Galvez, C., & Prairial, C. (1993). Effect of disorientation on visual analysis, familiarity decision and semantic decision on faces. *British Journal of Psychology*, 84(4), 433–441. https://doi.org/10.1111/j.2044-8295.1993.tb02494.x

Burleigh, T. J., Schoenherr, J. R., & Lacroix, G. L. (2013). Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in Human Behavior*, 29(3), 759–771. https://doi.org/10.1016/j.chb.2012.11.021

Cigna, M.-H., Guay, J.-P., & Renaud, P. (2015). La reconnaissance émotionnelle faciale : Validation préliminaire de stimuli virtuels dynamiques et comparaison avec les Pictures of Facial Affect (POFA). *Criminologie*, 48(2), 237. https://doi.org/10.7202/1033845ar

Cosker, D., Krumhuber, E., & Hilton, A. (2010). Perception of linear and nonlinear motion properties using a FACS validated 3D facial model. In: *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization* (pp. 101–108). ACM

Crivelli, C., & Fridlund, A. J. (2018). Facial Displays Are Tools for Social Influence. *Trends in Cognitive Sciences*, 22(5), 388–399. https://doi.org/10.1016/j.tics.2018.02.006

Crivelli, C., Jarillo, S., & Fridlund, A. J. (2016). A Multidisciplinary Approach to Research in Small-Scale Societies : Studying Emotions and Facial Expressions in the Field. *Frontiers in Psychology*, 7. https://doi.org/10.3389/fpsyg.2016.01073

Dalibard, S., Magnenat-Talmann, N., & Thalmann, D. (2012). Anthropomorphism of artificial agents: a comparative survey of expressive design and motion of virtual Characters and Social Robots. In *Workshop on Autonomous Social Robots and Virtual Humans at the 25th Annual Conference on Computer Animation and Social Agents (CASA 2012)*

Darwin, C., Ekman, P., & Prodger, P. (1998). *The Expression of the Emotions in Man and Animals*. Oxford University Press

David, L., Samuel, M. P., Eduardo, Z. C., & García-Bermejo, J. G. (2014). Animation of Expressions in a Mechatronic Head. In: *M. A. Armada, A. Sanfeliu, & M. Ferre (Eds.), ROBOT2013: First Iberian Robotics Conference* (Vol. 253, pp. 15–26). https://doi.org/10.1007/978-3-319-03653-3_2

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, 47(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, 48(1), 1–12. https://doi.org/10.3758/s13428-015-0567-2

Dodich, A., Cerami, C., Canessa, N., Crespi, C., Marcone, A., Arpone, M., ... Cappa, S. F. (2014). Emotion recognition from facial expressions: a normative study of the Ekman 60-Faces Test in the Italian population. *Neurological Sciences*, 35(7), 1015–1021. https://doi.org/10.1007/s10072-014-1631-x.

Du, S., & Martinez, A. M. (2015). Compound facial expressions of emotion : From basic research to clinical applications. *Dialogues in Clinical Neuroscience*, 17(4), 443–455

Ekman, P. (1971). *Universals and cultural differences in facial expressions of emotion*. University of Nebraska Press

Ekman, P. (1976). *Pictures of facial affect*. Consulting Psychologists Press

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. https://doi.org/10.1080/02699939208411068

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

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system: the manual*. OCLC: 178927696. Research Nexus.

Ekman, P., & Oster, A. H. (1979). Facial Expressions of Emotion. *Annual Review of Psychology*, 30(1), 527–554. https://doi.org/10.1146/annurev.ps.30.020179.002523

Ferrey, A. E., Burleigh, T. J., & Fenske, M. J. (2015). Stimulus-category competition, inhibition, and affective devaluation: a novel account of the uncanny valley. *Frontiers in Psychology*, 6. https://doi.org/10.3389/fpsyg.2015.00249.

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition & Emotion*, 22(6), 1094–1118. https://doi.org/10.1080/02699930701626582

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

Happy, S. L., Patnaik, P., Routray, A., & Guha, R. (2017). The Indian Spontaneous Expression Database for Emotion Recognition. *IEEE Transactions on Affective Computing*, 8(1), 131–142. https://doi.org/10.1109/TAFFC.2015.2498174

Kähler, K., Haber, J., & Seidel, H.-P. (2001). Geometry-based muscle modeling for facial animation. In: *Graphics interface* (Vol. 2001, pp. 37–46)

Krippendorff, K. (2004). *Content analysis: an introduction to its methodology* (2nd ed). Sage.

Krippendorff, K. (2011). *Computing Krippendorff's Alpha-Reliability*, from https://repository.upenn.edu/asc_papers/43/

Krumhuber, E. G., Skora, L., Küster, D., & Fou, L. (2017). A Review of Dynamic Datasets for Facial Expression Research. *Emotion Review*, 9(3), 280-292. https://doi.org/10.1177/1754073916670022

Krumhuber, E. G., Tamarit, L., Roesch, E. B., & Scherer, K. R. (2012). FACSGen 2.0 animation software: Generating three-dimensional

FACS-valid facial expressions for emotion research. *Emotion, 12*(2), 351–363. https://doi.org/10.1037/a0026632

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion, 24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. (pp. 94–101). https://doi.org/10.1109/CVPRW.2010.5543262.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed). Lawrence Erlbaum Associates.

Magnenat-Thalmann, N., & Thalmann, D. (Éds.). (2004). *Handbook of virtual humans*. Wiley.

Mavadati, S.M., Mahoor, M.H., Bartlett, K., Trinh, P., & Cohn, J.F. (2013). DISFA: A Spontaneous Facial Action Intensity Database. *IEEE Transactions on Affective Computing, 4*, 151–160.

Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. *IEEE Transactions on Affective Computing, 1–1.* https://doi.org/10.1109/TAFFC.2017.2740923. arXiv: 1708.03985

Mumenthaler, C., & Sander, D. (2012). Social appraisal influences recognition of emotions. *Journal of Personality and Social Psychology, 102*(6), 1118-1135. https://doi.org/10.1037/a0026885

Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods, Instruments, & Computers, 36*(4), 634–638. https://doi.org/10.3758/BF03206544

Pandzic, I. S., & Forchheimer, R. (Eds.). (2002). *MPEG-4 facial animation: the standard, implementation, and applications.* J. Wiley.

Pinet, S., Zielinski, C., Mathôt, S., Dufau, S., Alario, F.-X., & Longcamp, M. (2017). Measuring sequences of keystrokes with jsPsych: Reliability of response times and interkeystroke intervals. *Behavior Research Methods, 49*(3), 1163–1176. https://doi.org/10.3758/s13428-016-0776-3

Reimers, S., & Stewart, N. (2015). Presentation and response timing accuracy in Adobe Flash and HTML5/JavaScript Web experiments. *Behavior Research Methods, 47*(2), 309–327. https://doi.org/10.3758/s13428-014-0471-1

Roesch, E. B., Sander, D., Mumenthaler, C., Kerzel, D., & Scherer, K. R. (2010). Psychophysics of emotion: The QUEST for Emotional Attention. *Journal of Vision, 10*(3), 4-4. https://doi.org/10.1167/10.3.4

Russell, J. A. (1993). Forced-choice response format in the study of facial expression. *Motivation and Emotion, 17*(1), 41–51. https://doi.org/10.1007/BF00995206

Ryan, K. F., & Gauthier, I. (2016). Gender differences in recognition of toy faces suggest a contribution of experience. *Vision Research, 129*, 69-76. https://doi.org/10.1016/j.visres.2016.10.003

Sayette, M. A., Cohn, J. F., Wertz, J. M., Perrott, M. A., & Parrott, D. J. (2001). A psychometric evaluation of the facial action coding system for assessing spontaneous expression. *Journal of Non-verbal Behavior, 25*(3), 167–185.

Sneddon, I., McRorie, M., McKeown, G., & Hanratty, J. (2012). The Belfast Induced Natural Emotion Database. *IEEE Transactions on Affective Computing, 3*(1), 32–41. https://doi.org/10.1109/T-AFFC.2011.26

Tcherkassof, A., Dupré, D., Meillon, B., Mandran, N., Dubois, M., & Adam, J.-M. (2013). DynEmo: A video database of natural facial expressions of emotions. *The International Journal of Multimedia & Its Applications, 5*(5), 61–80.

Valstar, M. F., & Pantic, M. (2010). *Induced Disgust, Happiness and Surprise: an Addition to the MMI Facial Expression Database, 6*

Villagrasa, S., & Susín Sánchez, A. (2009). Face! 3d facial animation system based on facs. In: *IV Iberoamerican Symposium in Computer Graphics* (pp. 203–209)

Wexler, D. (1972). Methods for utilizing protocols of descriptions of emotional states. *Journal of Supplemental Abstract Services, 2*, 166

Yamada, Y., Kawabe, T., & Ihaya, K. (2013). Categorization difficulty is associated with negative evaluation in the "uncanny valley" phenomenon. *Japanese Psychological Research, 55*(1), 20–32. https://doi.org/10.1111/j.1468-5884.2012.00538.x

Obtaining the Software The software can be downloaded at this address: https://github.com/montybot/FACSHuman/