FACSGen 2.0 Animation Software: Generating 3D FACS-Valid Facial Expressions for Emotion Research

Eva G. Krumhuber and Lucas Tamarit

University of Geneva

Etienne B. Roesch

University of Reading

Klaus R. Scherer

University of Geneva

Author Note

Eva G. Krumhuber, Lucas Tamarit, and Klaus Scherer, Swiss Center for Affective Sciences, University of Geneva; Etienne B. Roesch, Centre for Integrative Neuroscience and Neurodynamics, University of Reading.

Eva G. Krumhuber is now at the School of Humanities and Social Sciences, Jacobs University Bremen, Bremen, Germany.

FACSGen 2.0 is available for scientific purposes only. Interested groups should contact the website at http://www.affective-sciences.org/facsgen-2010 or Klaus Scherer, Swiss Center

for Affective Sciences, CISA - University of Geneva, 7 Rue des Battoirs, 1205 Geneva, Switzerland. Email: Klaus.Scherer@unige.ch

This research was supported by the Swiss National Science Foundation, Humaine and Gfk Association. The authors would like to thank the four FACS coders for their help with the validation in Experiment 1 and Flavie Martin for data collection in Experiment 2.

Correspondence concerning this article should be addressed to Eva Krumhuber, Research IV, Campus Ring 1, Jacobs University Bremen, D-28759 Bremen, Germany. E-mail: e.krumhuber@jacobs-university.de

Word count: 5987

Abstract

In this article we present FACSGen 2.0, new animation software for creating static and dynamic 3-dimensional facial expressions on the basis of the Facial Action Coding System (FACS; Ekman, Friesen, & Hager, 2002). FACSGen permits total control over the Action Units (AUs), which can be animated at all levels of intensity and applied alone or in combination to an infinite number of faces. In 2 studies, we tested the validity of the software for the AU appearance defined in the FACS manual and the conveyed emotionality of FACSGen expressions. In Experiment 1, 4 FACS-certified coders evaluated the complete set of 35 single AUs and 54 AU combinations for AU presence/absence, appearance quality, intensity, and asymmetry. In Experiment 2, lay participants performed a recognition task on emotional expressions created with FACSGen software and rated the similarity of expressions displayed by human and FACSGen faces. Results showed good to excellent classification levels for all AUs by the 4 FACS coders, suggesting that the AUs are valid exemplars of FACS specifications. Lay participants' recognition rates for 9 emotions were high and comparisons of human and FACSGen expressions were very similar. The findings demonstrate the effectiveness of the software in producing reliable and emotionally valid expressions and suggest its application in numerous scientific areas, including perception, emotion, and clinical and neuroscientific research.

Keywords: emotion, facial expression, Facial Action Coding System, FACSGen, animation

FACSGen 2.0 Animation Software: Generating 3D FACS-Valid Facial Expressions for Emotion Research

The use of facial expressive stimuli has contributed much to our knowledge of the perception and recognition of emotions. Over the last years, several databases of emotion-specific expressions (MSFDE: Beaupré & Hess, 2005; JACFEE: Biehl et al., 1997; POFA: Ekman & Friesen, 1976; KDEF: Goeleven, de Raedt, Leyman, & Verschuere, 2008; RaFD: Langner et al., 2010; UCDSEE: Tracy, Robins, & Schriber, 2009; ADFES: Van der Schalk, Hawk, & Fischer, 2009; GEMEP: Bänziger & Scherer, 2010; Bänziger, Mortillaro, & Scherer, 2011) have been developed with the aim of providing standardized sets of emotional displays. These sets commonly show between six and nine distinct emotions (i.e., anger, fear, happiness, and sadness) expressed by Caucasian faces or those of other races. In addition, different versions of gaze and head orientation are often available, allowing variations of several characteristics. Although such facial displays are representative exemplars of emotion expressions and achieve good recognition rates, control over the type and number of variables is limited. For example, facial expressions generally differ between posers in intensity and underlying facial actions even when performed at high levels of skill (see Scherer & Ellgring, 2007). Opportunity is also limited for manipulating combinations of features and general properties of the face (i.e., age, ethnicity, gender). Moreover, most databases consist of emotional expressions presented as still photographs. Given the importance of motion in expression perception (i.e., Ambadar, Schooler, & Cohn, 2005; Bould, Morris, & Wink, 2008), the capacity to produce dynamic facial stimuli that can be systematically varied and controlled without sacrificing overall validity is urgently needed. The present article introduces FACSGen 2.0, new animation software for creating realistic three-dimensional (3D) facial expressions, both static and dynamic, in experimental research.

FACSGen permits total control over the stimulus material and corresponding informational cues (i.e., facial appearance), including lighting and observer's vantage point. Facial stimuli can be parametrically manipulated according to the experimenter's needs, opening possibilities for the systematic testing of specific hypotheses. FACSGen 2.0 is built on top of FaceGen Modeller (2007), an existing commercial tool for creating an infinite number of facial identities of any age, gender, and ethnicity. Photorealistic skin texture can be mapped onto the face, thereby simulating a unique, human-like appearance. In addition, we included different texture layers (i.e., diffuse color, ambient occlusion, and gloss and normal maps), which are combined during the rendering stage to achieve the final appearance. Specifically, the application of normal maps enables the simulation of small-scale wrinkles, bumps, and crevices and represents an extension of the original FaceGen system. Whereas FaceGen provides only limited control over the manipulation of facial expressions and offers a small number of inbuilt emotional expressions, the new FACSGen animation software allows the creation of facial expressions on the basis of objective descriptors, as provided by the Facial Action Coding System (FACS; Ekman & Friesen, 1978; Ekman et al., 2002).

FACS describes all possible visually distinguishable facial movements in terms of Action Units (AUs). An AU lists the appearance changes (i.e., shape alterations, motion direction, wrinkles, bulges, etc.) occurring with the contraction of a facial muscle group that can be controlled independently from all other facial muscle groups. In total, FACS contains 58 such AUs, of which 44 are commonly used to describe most facial expressions of emotion. The advantage of FACS is that the constituent AUs of any expression are analyzed separately, and their intensity (3-point scale), time course (onset, apex, offset), and asymmetry (L, R) can be objectively determined. In a first version of the software, we implemented a preliminary set of 16 AUs (see Roesch et al., 2011, for validation data of the general FACSGen approach). To refine the AU appearance quality and wrinkle detail, we redesigned all AUs in FACSGen 2.0 in collaboration with a professional computer graphics company.[1] Moreover, a large number of additional AUs were sculpted from descriptions of facial surface changes at maximum contraction by the FACS manual. The modeling process was closely monitored and rechecked by a FACS-certified coder (E.K.), who requested several revisions per AU, until the defined appearance changes were satisfactorily addressed. On the whole, we implemented 35 AUs, consisting of all upper and lower face AUs (except AU28), including several head and eye movements.

In FACSGen 2.0, each AU is represented by a software slider that provides control over the magnitude of the morph target in a value range from 0 to 100% (see Appendix A). AUs can be activated alone or in combination to create complex expressions. The intensity levels can be precisely defined, allowing the creation of identical expressions with equivalent parameter settings. In addition, AUs at different intensities can be combined to form new composite expressions that can then be used as separate morph targets. This enables the user to generate almost any emotionally expressive or nonemotional-specific facial expression. Besides the static manipulation, a separate window allows the nonlinear manipulation of activation curves of single AUs and AU combinations, offering the representation of detailed dynamic time courses (see Appendix B). Specifically, the duration and form of AUs can be systematically varied for onset, apex, and offset phase. The time profile and intensity of each phase is changeable and allows for the sophisticated simulation of facial movements. FACSGen 2.0 outputs both static and dynamic expressions, with a multitude of parameter (size, viewpoint, texture resolution, and background color) and lighting (azimuth, elevation, and intensity) options, as specified by the user. An advantage of the software is that it does not require any prior technical knowledge or expertise for producing high-quality animations. With a limited amount of training, interested research groups from any discipline can use it on a standard working platform (Windows OS). FACSGen 2.0 is available for noncommercial use by qualified research groups.

Overview of Validation Studies

In the following sections, we present two studies that aimed to test the validity of the full version of the FACSGen software for the AU appearances defined in FACS and the emotional meaning conveyed by FACSGen-generated expressions. Experiment 1 focused on the evaluation of the AUs for FACS rules and examined whether the AUs as synthesized in FACSGen 2.0 correspond to the detailed description in the FACS manual. In this evaluation phase, only FACS-certified coders participated and scored the complete set of 35 AUs for AU presence/absence, appearance quality, intensity, and asymmetry. Furthermore, to validate the single AUs in combination, the coders scored 46 AU combinations[2] that are listed in the FACS manual, as well as 8 emotion-specific AU combinations. All single AUs and AU combinations were displayed by four stimulus faces, representing both genders and two ethnicities. No emotion-inferential evaluations were made by any of the FACS coders in the first study.

Experiment 2 focused on the perceived emotionality of FACSGen expressions and investigated whether these expressions convey similar emotional meaning as that of human expressions. In this evaluation phase, lay participants first performed a recognition task on the emotion-specific AU combinations and rated the perceived intensity and believability for nine emotions portrayed by photofit FACSGen faces. To manipulate the perceived emotional magnitude, we displayed all expressions at high and medium intensity. If the FACSGen expressions were sufficiently realistic, we would expect to find ratings of accuracy, intensity, and believability that were similar to those previously reported with facial expression databases. To provide a stringent test of the appearance quality of FACSGen expressions, we asked participants to further perform a comparison task in which they viewed emotional expressions displayed by human faces and photofit FACSGen faces side by side. If perceived resemblances were high, this would suggest that FACSGen expressions reproduce the emotional signaling value of human expressions in a satisfactory fashion.

Experiment 1

The aim of the first study was to provide an exhaustive FACS validation of the software for all upper and lower face AUs (except AU28) and AU combinations, including several head and eye movements described in the FACS manual (Ekman et al., 2002). In addition, we also validated prototypical AU combinations of several basic and social emotions.

Method

Stimulus material and design. In total, 35 single AUs and 54 AU combinations were subject to validation (see Appendix C). AU combinations consisted of 46 nonemotional and 8 emotion-specific combinations (anger, disgust, embarrassment, fear, happiness, pride, sadness, and surprise). The targeted expressions of basic emotions were based on prototypes defined by Ekman and colleagues (Ekman & Friesen, 1978; Ekman et al., 2002). For social emotions (embarrassment and pride), we relied on descriptions provided by Keltner (1995), Tracy and Robins (2008), and Van der Schalk et al. (2009).

White and Black faces were used as stimulus targets to test for the generalization of AU appearance across different ethnicities. Two White and Black male and female faces expressed all single AUs and AU combinations. For validation purposes, every AU was presented to each FACS coder in a different face. The representation of the four target faces was counter-balanced across the different AUs. To obtain measures of interrater reliability, FACS coders used the same face to code 12% of the stimulus material (six single AUs and AU combinations).

For every stimulus face, we generated video clips in which the single AU or AU combination linearly unfolded (onset duration = 1,000 ms) until reaching its peak (apex duration = 1,000 ms) and returning to a neutral baseline. Dynamic expressions were synthesized at a frame rate of 25 images per second and lasted a total of 3 s. In addition, static images that have been extracted from the video clips were used. For single AUs, these images showed the AU at three different levels of intensity of the morph target: 30% (low), 60% (medium), and 90% (high). The intensity levels were chosen in such a way as to correspond as closely as possible to the 3-point intensity scoring in FACS (x, y, z). For AU combinations, static images showed the expression at the peak level of morph targets with 70% magnitude (nonemotional) or varying magnitude (emotion-specific) of the AUs. All video clips and images were rendered in color with the same viewpoint, camera focal length, and lighting. The resulting set of stimuli measured 600 × 1,000 pixels and was displayed on a black background in random order. Video examples of the single AUs (Video 1) and AU combinations (Video 2) can be viewed at http://www.affective-sciences.org/facsgen-2010.

Procedure. Four certified FACS coders participated in the validation phase. Each coder received two coding sets. The first set contained video clips of single AUs and pictures that showed the AU at 30%, 60%, and 90% intensity. The second set contained video clips of AU combinations and pictures of the AU combinations at the peak of the expression. In addition, neutral pictures of the four stimulus faces were provided. FACS coders were free to watch all video clips and pictures before they started with the scoring. However, they were blind to the type and number of AUs that were part of an expression in both coding sets. Overall, the FACS scoring procedure required about 12-15 hr of work per coder.

Dependent measures. For the first coding set of single AUs, FACS coders were instructed to score: (a) the presence of the AU; (b) AU asymmetry (if applicable); (c) AU quality on a 7-point scale ("How well does the AU match the appearance changes described in the FACS manual?", ranging from 1 [*very poor match*] to 7 [*very good match*]); and (d) AU intensity on a 3-point scale (x, y, z) of the 30%, 60%, and 90% pictures. For the second coding set of AU combinations, FACS coders had to score (a) the presence of the AU and (b) AU asymmetry (if applicable). No intensity ratings were made for AU combinations.

Results and Discussion

For all single AUs and AU combinations, we calculated the number of cases in which the scoring of the four FACS coders corresponded to the target AU formula. If the coding deviated from the target formula (i.e., by coding an additional AU or failing to code a target AU), it was counted as incorrect. Note that this high degree of required accuracy constituted an extremely stringent test for AU validity (including AU combinations). Table 1 shows the mean classification and interrater reliability results of the 35 single AUs and 54 AU combinations.

Overall, the validation data showed good to excellent classification results for all AUs. Ninety-eight percent of all single AUs matched the target AU formula. Except in two cases in which one of the four FACS coders provided an AU score that was different from that of the target formula, all AUs were coded accurately. For all single AUs, quality of AU appearance was scored highly (M = 6.37, SE = 0.06) and classification results of AU intensity showed sufficient accuracy at the three levels of intensity. The proposed 30-intensity, 60-intensity, and 90-intensity level can therefore be used in accordance with the FACS specifications of x-intensity (low), y-intensity (medium), and z-intensity (high), respectively. Interrater agreement between the four FACS coders was good to excellent for all single AUs, including intensity (intraclass correlations ranging from 0.79 to 0.99). The same pattern of results was evident for the reliability items in which the four FACS coders scored six AUs for the same face. Only for the 90-intensity reliability coding was classification success lower (75%). However, in none of the cases did more than two FACS coders suggest an intensity level that was different from the target intensity.

For the 54 AU combinations, classification accuracy was similarly high at 80%, with excellent interrater agreement. There were no overall differences in accuracy between the nonemotional and emotion-specific AU combinations. In most cases, only one AU of the AU combination deviated from the target formula or was omitted from coding. For example, AU26 (Jaw Drop) instead of AU27 (Mouth Stretch) was scored by one of the four FACS coders for surprise, whereas AU5 (Upper Lid Raiser) was left out by one coder for anger and fear. Besides these minor deviations of singular AUs from the target formula, all FACS coders agreed on the majority of AUs in each AU combination, which is reflected in the high interrater agreement (ranging from 0.95 to 1.00). The pattern of results was the same for the six reliability items in which the four FACS coders scored six AU combinations for the same face. Eighty-three percent of reliability AU combinations were accurately coded, and interrater agreement was high at 0.95.

Experiment 2

The high accuracy of classification for all single AUs and AU combinations suggests that the validation of AUs as synthesized by FACSGen 2.0 was successful. Consequently, all AUs achieved verification by FACS-certified coders for the relevant target AU formula. In the second experiment, we aimed to test the validity of FACSGen expressions for emotional meaning. For this purpose, we focused on the emotion-specific AU combinations of the first study (including contempt) and obtained participants' emotion recognition scores as well as their ratings of intensity and believability. Furthermore, we conducted a comparison task in which participants judged the similarity of emotional expressions displayed by FACSGen faces and human faces.

Method

Participants. Thirty-nine students (34 women, 5 men) from the University of Geneva participated in exchange for course credit or CHF15. Their mean age was 22.9 years (*SD* = 3.54), ranging from 18 to 38 years.

Stimulus material and design.

Recognition task. Two White male and female photofit FACSGen faces were used as stimulus targets. Photofit faces contribute to a realistic facial appearance by integrating the texture detail of a real human face, such as facial hair (e.g., eyebrows) and skin pigmentation. All photofit FACSGen faces expressed the eight emotion-specific AU combinations (anger, disgust, embarrassment, fear, happiness, pride, sadness, and surprise) that had been validated in the previous study. In addition, we included contempt, which was operationalized as a unilateral dimpler (AU14uni) from descriptions by Langner et al. (2010). To manipulate the degree of perceived emotional magnitude, we displayed expressions at two intensity levels (100%-high, 50%-medium). Figure 1 shows examples of each emotion as expressed by a photofit FACSGen face at high intensity.

For every stimulus face, dynamic emotional expressions were created at a frame rate of 25 frames per second. Stimuli started at a neutral position and then changed linearly (onset duration = 1,500 ms) to a peak expression with an apex duration of 1,500 ms. In total, the video clips for each emotion covered 3 s. The four photofit faces showing nine different emotional expressions at two intensity levels were rendered in color with the same viewpoint, camera focal length, and lighting. The resulting set of 72 stimuli measured 800 × 1,200 pixels and was displayed on a black background in random order. Video examples of each type of emotional expression at 100-intensity (Video 3) and 50-intensity (Video 4) can be viewed at http://www.affective-sciences.org/facsgen-2010.

Comparison task. Pictures of four human faces (two male, two female) were selected from the Amsterdam Dynamic Facial Expressions Set (ADFES; Van der Schalk et al., 2009) and showed the nine emotions (including neutral) at peak level. All expressions were validated in FACS terms and achieved sufficient emotion recognition rates. Based on the detailed FACS coding of these human expressions, photofit expressions of the same faces were modeled in FACSGen. That is, the same AUs as coded in the human expressions were implemented in photofit FACSGen expressions. As we were unable to resynthesize the human hairstyle, oval masks were used to conceal the outer part of the face. Human and FACSGen expressions always showed the same emotions and appeared side by side (with the presentation side being counterbalanced). The resulting set of 40 stimuli (4 faces × 10 emotions) measured 834 × 569 pixels and was displayed on a black background in random order for 5 s each. Figure 2 shows examples of each emotion as displayed by human faces and photofit FACSGen faces.

Procedure. After signing a consent form, participants received detailed instructions regarding the purpose of the study and the experimental tasks with Eprime 2.0.8.22 (Psychology Software Tools, Inc.). The experiment always started with the recognition task in which dynamic expressions of nine emotions were shown by four photofit FACSGen faces at high and medium intensity. Participants were informed that they would see short video clips of animated characters displaying various facial expressions. Their task was to indicate which emotion was being expressed in the face, and how intense and believable the facial expression was in terms of the chosen emotion category. For the comparison task, paired images of 10 emotions were shown by real human faces and FACSGen faces next to each other. Participants were told that the (FACSGen) computer-generated expressions were always modeled on the corresponding real human facial expression by the person shown. No information was provided about the type of emotion expressed by the human face, so that comparisons had to rely on actual resemblance of expressive features. The participants' task was to indicate how well the respective expression shown by the human person was captured in the FACSGen animation.

Dependent measures. In the recognition task, participants successively rated for every stimulus (a) the expressed emotion, (b) the intensity, and (c) the believability of the expression in terms of the chosen emotion category. In line with previous research (e.g., Biehl et al., 1997; Goeleven et al., 2008; Langner et al., 2010), expressed emotion was measured within a fixed-choice format that required the selection of an emotion category that best matched the shown facial expression. Response categories included the nine presented emotions, as well as the option "no emotion/other emotion" if none of the suggested categories was considered applicable (Frank & Stennett, 2001). For the intensity and believability assessment of the chosen emotion, ratings were made on 7-point Likert scales, with response options ranging from 1 (*not at all*) to 7 (*very*).

In the comparison task, participants indicated for each image pair how well the (FACSGen) computer-generated expression captured and reproduced the human expression. Response options ranged from 1 (*not well at all*) to 7 (*very well*).

Results and Discussion

Recognition accuracy. Analyses of variance (ANOVAs) with the within-subjects factors emotion (anger, contempt, disgust, embarrassment, fear, happiness, pride, sadness, and surprise) and intensity (100, 50) were conducted on the recognition scores. Table 2 shows the mean percentage recognition and unbiased hit rates for the nine emotions at two intensity levels. Percentage recognition refers to the percentage of correctly identified expressions and was calculated as the number of correct responses divided by the number of target stimuli for an emotion. As this measure does not take response bias into account (e.g., the bias to say "happy" for all expressions), we also calculated unbiased hit rates (Wagner, 1993). Unbiased hit rates express recognition accuracy as proportions of both stimulus frequency and response frequency and vary between 0 and 1 (perfect recognition; see Goeleven et al., 2008, for a detailed description of unbiased hit rates).

The mean overall percentage recognition for emotions was 72%. Recognition rates were sufficiently high at 100-intensity and 50-intensity and comparable to those reported in previous research with human faces (Bänziger, Grandjean, & Scherer, 2009; Bänziger et al., 2011; Bänziger & Scherer, 2010; Beaupré & Hess, 2005; Goeleven et al., 2008; Langner et al., 2010; Tracy et al., 2009; Van der Schalk et al., 2009). Except for contempt at 50-intensity ($M = 0.40$, $p = .200$), all percentage recognition scores and unbiased hit rates were significantly higher than chance, set conservatively at 33%, *p*s < .01 (Tracy et al., 2009). An ANOVA of the arcsine-transformed unbiased hit rates (Winer, 1971) revealed a significant main effect of intensity, $F(1, 38) = 30.09$, $p < .001$, $\eta_p^2 = .44$. Overall, expressions displayed at 100-intensity ($M = 0.68$, $SE = .02$) were better recognized than those displayed at 50-intensity ($M = 0.59$, $SE = .03$). In addition, there was a significant main effect of emotion, $F(8, 304) = 7.19$, $p < .001$, $\eta_p^2 = .16$. Recognition rates were significantly higher for surprise, anger, and sadness ($M = 0.74$, $SE = .04$)

and significantly lower for contempt (*M* = 0.45, *SE* = .04), compared with all other expressions (means between 0.56 and 0.67, *p*s < .05). The low recognition of contempt replicates the findings of Langner et al. (2010) and Van der Schalk et al. (2009), who also found contempt to be the least well-recognized expression. For fear and embarrassment, similar suboptimal recognition results were reported by Beaupré and Hess (2005), Goeleven et al. (2008), and Tracy et al. (2009). There was no significant interaction between intensity and emotion, $F(8, 304) = 1.07$, $p = .38$, $\eta_p^2 = .03$.

Intensity. For intensity ratings, a 9 (emotion) × 2 (intensity) ANOVA revealed a significant main effect of target intensity, $F(1, 38) = 179.67$, $p < .001$, $\eta_p^2 = .82$. As expected, expressions at 100-intensity (*M* = 5.09, *SE* = .11) were judged as being more intense than those at 50-intensity (*M* = 3.99, *SE* = .13), confirming that the manipulation of intensity of the emotional expressions was successful. Furthermore, a significant main effect of emotion occurred, $F(8, 304) = 22.27$, $p < .001$, $\eta_p^2 = .37$. Overall, surprise and pride were rated to be the most intense expressions (*M* = 5.01, *SE* = .13), followed by anger and fear; then happiness, disgust, and sadness (means between 4.78 and 4.46); and finally embarrassment and contempt (*M* = 4.00, *SE* = .14). These findings are in line with previous results in which intensity ratings were among the highest for surprise and the lowest for contempt (Goeleven et al., 2008; Langner et al., 2010). The main effects of intensity and emotion were qualified by a significant two-way interaction between intensity and emotion, $F(8, 304) = 4.41$, $p < .001$, $\eta_p^2 = .10$. Depending on the level of target intensity, emotions differed significantly from each other in their ratings of intensity (see Table 2). Post hoc tests showed that judged intensity of anger and happiness varied considerably across the 100-intensity and 50-intensity condition. Whereas anger at 100-intensity was rated as being more intense than happiness, fear, disgust, and sadness (*p*s < .05), these differences dropped to insignificance at 50-intensity (*p*s > .05). Similarly, intensity ratings of happiness that differed from those of contempt and embarrassment at 100-intensity (*p*s < .001) were not significantly different at 50-intensity (*p*s > .05). In this sense, ratings of intensity tended to merge with lower target intensity of the emotional expressions.

Believability. A 9 (emotion) × 2 (intensity) ANOVA on the believability ratings showed a significant main effect of intensity, $F(1, 38) = 10.29$, $p < .01$, $\eta_p^2 = .21$. Overall, expressions displayed at 100-intensity ($M = 4.82$, $SE = .16$) were rated to be more believable than expressions displayed at 50-intensity ($M = 4.49$, $SE = .14$). A significant main effect of emotion revealed significant differences in perceived believability between the emotions, $F(8, 304) = 9.75$, $p < .001$, $\eta_p^2 = .20$. In general, surprise, pride, happiness, and anger were rated to be most believable ($M = 4.98$, $SE = .17$), followed by sadness and embarrassment ($M = 4.62$, $SE = .16$), with contempt, fear, and disgust scoring around the midpoint of the scale ($M = 4.26$, $SE = .19$). This pattern of results was highly similar for expressions at 100-intensity and 50-intensity (see Table 2) and comparable to genuineness ratings reported by Langner et al. (2010) for human expressions. The interaction between intensity and emotion was not significant, $F(8, 304) = 1.49$, $p = .16$, $\eta_p^2 = .04$.

Comparison of human and FACSGen expressions. To examine how closely participants rated emotional expressions displayed by human and FACSGen faces, we computed a one-way ANOVA with the within-subjects factor emotion (anger, contempt, disgust, embarrassment, fear, happiness, pride, sadness, surprise, and neutral) on the similarity measure. Results showed that the main effect of emotion was significant, $F(9, 342) = 8.46$, $p < .001$, $\eta_p^2 = .18$. As expected, for neutral expressions, FACSGen faces were judged to be most like human faces ($M = 5.57$, $SE = .16$), which corresponds to a similarity measure of 80% (see Figure 3).

Note that these expressions showed only neutral photofit faces that had undergone no emotional manipulation.[3] The result of the neutral expressions can therefore function as a baseline for the interpretation of the emotional expressions. Overall, mean similarity across all emotions was 4.85 (SE = .17), thereby demonstrating high comparability in expressive quality. Surprise and anger were rated to be most similar (M = 5.28, SE = .15) between FACSGen and human faces, followed by contempt, sadness, happiness, fear, and disgust (M = 4.83, SE = .17), and finally by embarrassment and pride (M = 4.46, SE = .18, ps < .05).

General Discussion

In this paper, we presented FACSGen 2.0, new animation software providing high-quality 3D facial stimuli for use in emotion expression research. FACSGen allows the creation of realistic facial expressions, both static and dynamic, on the basis of FACS. Facial stimuli and related informational cues can be parametrically controlled and manipulated for a virtually infinite number of faces, allowing the production of standardized stimulus material for experimental research. In two studies, we tested the validity of the software for the AU appearance defined in FACS and the emotional meaning conveyed by FACSGen expressions. Experiment 1 reported validation data for 35 single AUs and 54 AU combinations that had been implemented in faces of different gender and ethnicity. The classification of AUs was high and the AUs interacted predictably in combination with each other. For all AUs, quality of AU appearance was scored satisfactorily by the FACS coders, and the three-level intensity coding generally matched the FACS specifications. Based on the high classification rates in combination with the good interrater reliabilities, these results suggest that the AUs as synthesized by FACSGen 2.0 are valid exemplars that correspond to what is described in the FACS manual.

Experiment 2 showed that emotional expressions generated with FACSGen convey affective meaning that is reliably recognized by lay participants. The mean recognition rate of 72% was high and comparable with those previously reported with human faces (Beaupré & Hess, 2005; Goeleven et al., 2008; Langner et al., 2010; Tracy et al., 2009; Van der Schalk et al., 2009). Overall, surprise, anger, and sadness were the most easily recognizable emotions, whereas expressions of contempt were most difficult to detect. The low recognition rate of contempt was in line with findings by Langner et al. (2010) and Van der Schalk et al. (2009), who argued that this may be a general feature of the emotion, and not of the expression itself. The manipulation of the perceived emotional magnitude was successful, with greater levels of intensity being attributed to expressions of full intensity than to expressions of medium intensity. Such high-intensity expressions were also better recognized and judged to be more believable than medium-intensity expressions, probably because of their increased emotional salience. When comparing emotional expressions displayed by FACSGen faces and human faces side by side, perceived resemblances were high. Similarity ratings for all nine emotions were significantly above the midpoint of the scale, suggesting that the emotional signal value of human expressions is sufficiently reproduced in FACSGen expressions. These findings underscore the effectiveness of the software in eliciting reliable and prototypical affective stimuli that can be used for systematic testing in emotion research.

FACSGen 2.0 is comparable to other software such as Poser (Spencer-Smith et al., 2001), FACE (Wehrle, Kaiser, Schmidt, & Scherer, 2000), realEmotion (Grammer, Tessarek, & Hofer, 2011), Alfred (Bee, Falk, & André, 2009), or the Virtual Actor Project (Helzle, Biehn, Schlömer, & Linner, 2004). Although some of these programs allow one to generate AU-based facial actions, we are unaware of whether and how they have been validated in FACS terms. Apart from the fact that these animation packages are not easily available or have become obsolete, there is a great difference in the ease of use, often requiring prior technical knowledge. FACSGen has the advantage of striking a balance between usability and realism. Currently, it is the only software to include FACS-validated AUs that can be used by researchers of any discipline. No special training is required to generate high-quality facial animations, and experimental stimuli can be produced rapidly. The software allows the creation of facial actions in FACS-defined form and motion at all levels of intensity. Moreover, the user has maximum flexibility by combining AUs of any type and by specifying the time profiles of facial movements in a linear or nonlinear fashion. To our knowledge, none of these options are offered in any other software currently available.

Despite these advantages of FACSGen 2.0 software, some limitations should be acknowledged. Because FACSGen models come by default with no hair, the appearance of faces, in particular female faces, may seem somewhat unusual. However, no restrictions are implied in the manifestation of gender typicality (Freeman & Ambady, 2009). Past research has shown that hairless synthetic faces are unambiguously recognized as male or female (Becker, Kenrick, Neuberg, Blackwell, & Smith, 2007; Roesch et al., 2011) because of variations in facial features. Moreover, trait attributions of male and female synthetic faces without hair have been found to be similarly sensitive to features resembling emotional expressions, as is the case for human faces with hair (see Becker et al., 2007; Oosterhof & Todorov, 2008). Inferences drawn from hairless faces, therefore, may not necessarily be separate from those drawn from faces with hair. Nonetheless, to address this issue, we are currently working on an automatic masking system that will conceal the peripheral part of the head, including the hair. A similar approach has been taken by Goeleven et al. (2008), who removed the hairline from the faces in their human database, arguing that this makes the emotional expression even more distinctive. We believe that our solution is a reasonable compromise, but acknowledge the possible limitations that may be caused by lack of hair.

Photofitting now allows the application of texture details such as facial hair (e.g., eyebrows, beard) and skin pigmentation to FACSGen models. Although this represents a significant advance in the human-like appearance of faces, miscellaneous components such as glasses, earrings, or other aesthetic items (i.e., piercing) cannot yet be included. We also observed that the sclera and the teeth are perceived as being too white, particularly in the comparison between FACSGen and human faces of Experiment 2. We have taken note of these issues and plan to correct for the brightness level in the future.

Stimuli created with FACSGen are 2D projections of actual 3D content. This 3D content can be rendered from any viewpoint, potentially allowing the presentation of stimuli in stereoscopic immersive environments of any kind. Like other software (e.g., Poser, Studio Max), however, the current implementation does not include such stereoscopic output. Moreover, a linear morph model is used to synthesize geometric movement of facial actions. Although this may not allow for an exact representation of naturally deforming motion of nonlinear quality, such linear blend shape approaches are still commonly used with high success in computer graphics (see Oleg, Rogers, Lambeth, Chiang, & Debevec, 2009; Parke & Waters, 1996).[4] Currently, FACSGen permits the generation of linear and nonlinear temporal motion by manipulating the activation curves of AUs. This provides the opportunity to resynthesize sophisticated facial behavior in a 3D and dynamic form. Until now, most human databases have consisted of static photographs of facial expressions, despite their rather unrealistic nature. This is surprising, given the large body of evidence showing that dynamic expressions are perceived as more naturalistic, realistic, and intense and that they evoke stronger facial and brain activation than do static expressions (Biele & Grabowska, 2006; Sato, Fujimura, & Suzuki, 2008; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Sato & Yoshikawa, 2007; Weyers, Mühlberger, Hefele, & Pauli, 2006).

FACSGen has already been valuable in several psychological and neuroscientific studies involving dynamic stimuli (Cristinzio, N'Diaye, Seeck, Vuilleumier, & Sander, 2010; N'Diaye, Sander, & Vuilleumier, 2009). Moreover, it could be a useful tool in the context of clinical applications requiring the training and rehabilitation of patients with emotional dysfunctions and facial movement disorders (see Denlinger, VanSwearingen, Cohn, & Schmidt, 2008). Because single facial actions can be activated dynamically and independently from each other, FACSGen allows the dissection of complex facial expressions into its parts. Acquiring such controlled facial stimuli of human posers remains a challenging and labor-intensive task. With FACSGen, these problems can be overcome, as facial expressions can be systematically deformed and controlled on the basis of objective descriptors, such as the Facial Action Coding System.

References

Ambadar, Z., Schooler, J., & Cohn, J. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science, 16,* 403–410.

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,* 691–704.

Bänziger, T., Mortillaro, M., & Scherer, K. R. (2011). *Introducing a new multimodal expression corpus for experimental research on emotion perception.* Manuscript submitted for publication.

Bänziger, T., & Scherer, K. R. (2010). Introducing the Geneva Multimodal Emotion Portrayal (GEMEP) corpus. In K. R. Scherer, T. Bänziger, & E. B. Roesch (Eds.), *Blueprint for affective computing: A sourcebook* (pp. 271–294). Oxford, England: Oxford University Press.

Beaupré, M. G., & Hess, U. (2005). Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology*, *36*, 355-370.

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, *92*, 179-190.

Bee, N., Falk, B., & André, E. (2009). Simplified facial animation control utilizing novel input devices: A comparative study. *Proceedings of the 13th International Conference on Intelligent User Interfaces* (IUI '09) (pp. 197–206). Sanibel Island, FL: Association for Computing Machinery.

Biele, C., & Grabowska, A. (2006). Sex differences in perception of emotion intensity in dynamic and static facial expressions. *Experimental Brain Research, 26*, 1-6.

Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., & Ton, V. (1997). Matsumoto and Ekman's Japanese and Caucasian facial expressions of emotion (JACFEE): Reliability data and cross-national differences. *Journal of Nonverbal Behavior, 21*, 3-21.

Bould, E., Morris, N., & Wink, B. (2008). Recognising subtle emotional expressions: The role of facial movements. *Cognition and Emotion, 22*, 1569-1587.

Cosker, C., Krumhuber, E., & Hilton, A. (2010). A FACS validated 3D human facial model. *Proceedings of the ACM/SSPNet 2nd International Symposium on Facial Analysis and Animation (FAA)* (p. 10). Edinburgh, UK.

Cristinzio, C., N'Diaye, K., Seeck, M., Vuilleumier, P., & Sander, D. (2010). Integration of gaze direction and facial expression in patients with unilateral amygdale damage. *Brain, 133*, 248–261.

Denlinger, R. L., VanSwearingen, J. M., Cohn, J. F, & Schmidt, K. L. (2008). Puckering and blowing facial expressions for patients with facial movement disorders. *Physical Therapy Journal, 88*, 909-915.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *The Facial Action Coding System* (2nd ed.). Salt Lake City, UT: Research Nexus eBook.

FaceGen Modeller (Version 3.2) [Computer software]. (2007). Retrieved from http://www.facegen.com/

Frank, M. G., & Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology, 80*, 75–85.

Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science*, *20*, 1183-1188.

Goeleven, E., de Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinksa Directed Emotional Faces: A validation study. *Cognition and Emotion*, *22*, 1094–1118.

Grammer, K., Tessarek, A., & Hofer, G. (2011). From emoticons to avatars: The simulation of facial expression. In A. Kappas & N. Krämer (Eds.), *Face-to-face communication over the Internet* (pp. 237–279). Cambridge, UK: Cambridge University Press.

Helzle, V., Biehn, C., Schlömer, T., & Linner, F. (2004). Adaptable setup for performance driven facial animation. *Proceedings of the International Conference on Computer Graphics and Interactive Techniques* (SIGGRAPH) (p. 54). Los Angeles, CA: Association for Computing Machinery.

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.

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 and Emotion*, *24*, 1377–1388.

N'Diaye, K., Sander, D., & Vuilleumier, P. (2009). Self-relevance processing in the human amygdala: Gaze direction, facial expression, and emotion intensity. *Emotion, 9*, 798–806.

Oleg, A., Rogers, M., Lambeth, W., Chiang, M., & Debevec, P. (2009). The Digital Emily project: Photoreal facial modeling and animation. *Proceedings of the International Conference on Computer Graphics and Interactive Techniques* (ACM SIGGRAPH) (pp.1-15), New York, NY: Association for Computing Machinery.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of facial evaluation. *Proceedings of the National Academy of Sciences, 105*, 11087-11092.

Parke, F. I., & Waters, K. (1996). *Computer facial animation*. Wellesley, MA: A. K. Peters Ltd.

Roesch, E. B., Tamarit, L., Reveret, L., Grandjean, D., Sander, D., & Scherer K. R. (2011). FACSGen: A tool to synthesize realistic, static and dynamic emotional facial expressions based on facial action units. *Journal of Nonverbal Behavior, 35*, 1-16.

Sato, W., Fujimura, T., & Suzuki, N. (2008). Enhanced facial EMG activity in response to dynamic facial expressions. *International Journal of Psychophysiology, 70*, 70-74.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Cognitive Brain Research, 20*, 81-91.

Sato, W., & Yoshikawa, S. (2007). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition, 104*, 1-18.

Scherer, K. R., & Ellgring, H. (2007). Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion, 7*, 113-130.

Spencer-Smith, J., Wild, H., Innes-Ker, A. H., Townsend, J., Duffy, C., Edwards, C., . . . Paik, J. W. (2001). Making faces: Creating three dimensional parameterized models of facial expression. *Behavior Research Methods, Instruments, & Computers, 33*, 115-123.

Tracy, J. L., & Robins, R. W. (2008). The automaticity of emotion recognition. *Emotion, 8*, 81-95.

Tracy, J. L., Robins, R. W., & Schriber, R. A. (2009). Development of a FACS-verified set of basic and self-conscious emotion expressions. *Emotion*, *9*, 554-559.

Van der Schalk, J., Hawk, S. T., & Fischer, A. H. (2009, July). *Validation of the Amsterdam Dynamic Facial Expression Set (ADFES)*. Poster presented at the meeting of the International Society for Research on Emotions (ISRE), Leuven, Belgium.

Wagner, H. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, *17*, 3-28.

Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality & Social Psychology*, *78,* 105-119.

Weyers, P., Mühlberger, A., Hefele, C., & Pauli, P. (2006). Electromyographic responses to static and dynamic avatar emotional facial expressions. *Psychophysiology*, *43*, 1-4.

Winer, B. (1971). *Statistical principles in experimental design* (2nd ed.). New York, NY: McGraw-Hill.

Footnotes

[1]Trait d'Esprit, http://www.traitdesprit.ch

[2]In cases in which an AU combination was not the sole aggregation of two or more individual AUs, but revealed new appearance changes, separate morph targets were created. To our knowledge, this was required only for the combinations AU1+4 and AU1+2+4.

[3]If comparisons between human and FACSGen expressions had been made simply on the basis of emotion categorization (thereby generalizing over a wide range of variants of an emotional expression), we would expect correspondence ratings to be considerably higher for neutral expressions (achieving ceiling rates close to 100%). Given that the perceived similarity of the two types of stimuli was not perfect even for neutral expressions, participants indeed seemed to rely on feature resemblance over and above whether the two expressions were recognizable as members of the same class of emotion.

[4]Clearly, more information should be gained in the future about the dynamics of facial actions through the quantitative analysis of facial movements over time in a variety of communicative contexts. It must be noted, however, that such real-time AU movements can be captured only with the use of dynamic 3D facial scanners or optical motion capture systems with a large number of markers, thereby allowing a comparison between linear and nonlinear geometric motion. Although efforts have recently begun to build a FACS-valid facial model based on nonlinear geometric movements recorded from real faces (see Cosker, Krumhuber, & Hilton, 2010), it will take several more years until such models become available for wider public distribution.

Table 1

Mean Correct Classification and Interrater Reliability for 35 Single Action Units (AUs) and 54 AU Combinations

FACS Coding Set	% Correct	Intraclass
(a) 35 Single AUs		
AUs	98.57	0.99
30-Intensity (x)	97.86	0.85
60-Intensity (y)	86.43	0.79
90-Intensity (z)	90.00	0.93
6 Reliability AUs		
AUs	100.00	1.00
30-Intensity (x)	100.00	1.00
60-Intensity (y)	87.50	0.96
90-Intensity (z)	75.00	0.97
(b) 54 AU combinations	80.09	0.97
46 nonemotional	82.87	0.96
8 emotion specific	81.25	0.98
6 Reliability AU combinations	83.33	0.95

Table 2

Means (Standard Errors) of Percentage Recognition, Unbiased Hit Rates, and Intensity and Believability Ratings as a Function of Emotion and Intensity

	Emotion								
Measure	Anger	Contempt	Disgust	Embarrassment	Fear	Happiness	Pride	Sadness	Surprise
100-Intensity									
% Recognition	87.82	56.41	68.59	69.23	72.44	88.46	74.36	83.97	87.82
	(4.09)	(4.57)	(5.09)	(5.84)	(4.19)	(2.88)	(4.73)	(4.45)	(3.66)
Unb. hit rate	0.82_{a}	0.50_{c}	0.61_{bc}	0.59_{c}	0.65_{bc}	0.72_{ab}	0.63_{bc}	0.80_{a}	0.81_{a}
	(.04)	(.05)	(.05)	(.06)	(.04)	(.04)	(.05)	(.05)	(.04)
Intensity	5.54_{a}	4.22_{e}	5.08_{c}	4.65_{d}	5.09_{c}	5.23_{bc}	5.47_{ab}	4.97_{c}	5.56_{a}
	(.13)	(.17)	(.16)	(.10)	(.13)	(.16)	(.14)	(.13)	(.12)
Believability	5.08_{ab}	4.34_{d}	4.26_{cd}	4.80_{bc}	4.51_{cd}	5.21_{a}	5.19_{a}	4.79_{bc}	5.24_{a}
	(.21)	(.20)	(.28)	(.14)	(.25)	(.20)	(.18)	(.22)	(.18)

50-Intensity									
% Recognition	71.79	48.08	61.54	60.26	67.31	77.56	71.79	76.28	87.82
	(4.79)	(5.31)	(5.87)	(5.09)	(4.12)	(5.10)	(5.45)	(4.40)	(3.66)
Unb. hit rate	0.66_{ab}	0.40_{d}	0.55_{ab}	0.53_{b}	0.57_{bc}	0.61_{ab}	0.59_{ab}	0.66_{ac}	0.73_{a}
	(.05)	(.05)	(.06)	(.05)	(.04)	(.05)	(.05)	(.05)	(.04)
Intensity	4.03_{b}	3.53_{c}	3.96_{b}	3.61_{c}	4.06_{b}	3.81_{bc}	4.43_{a}	3.95_{b}	4.56_{a}
	(.15)	(.19)	(.16)	(.16)	(.14)	(.19)	(.16)	(.15)	(.14)
Believability	4.65_{a}	4.34_{bc}	4.01_{c}	4.27_{c}	4.09_{c}	4.72_{a}	4.90_{a}	4.61_{ab}	4.86_{a}
	(.18)	(.17)	(.20)	(.18)	(.19)	(.18)	(.16)	(.18)	(.18)

Note. Means in the same row not sharing a subscript differ significantly. For reasons of readability, unbiased (unb.) hit rates are reported here as untransformed proportions.

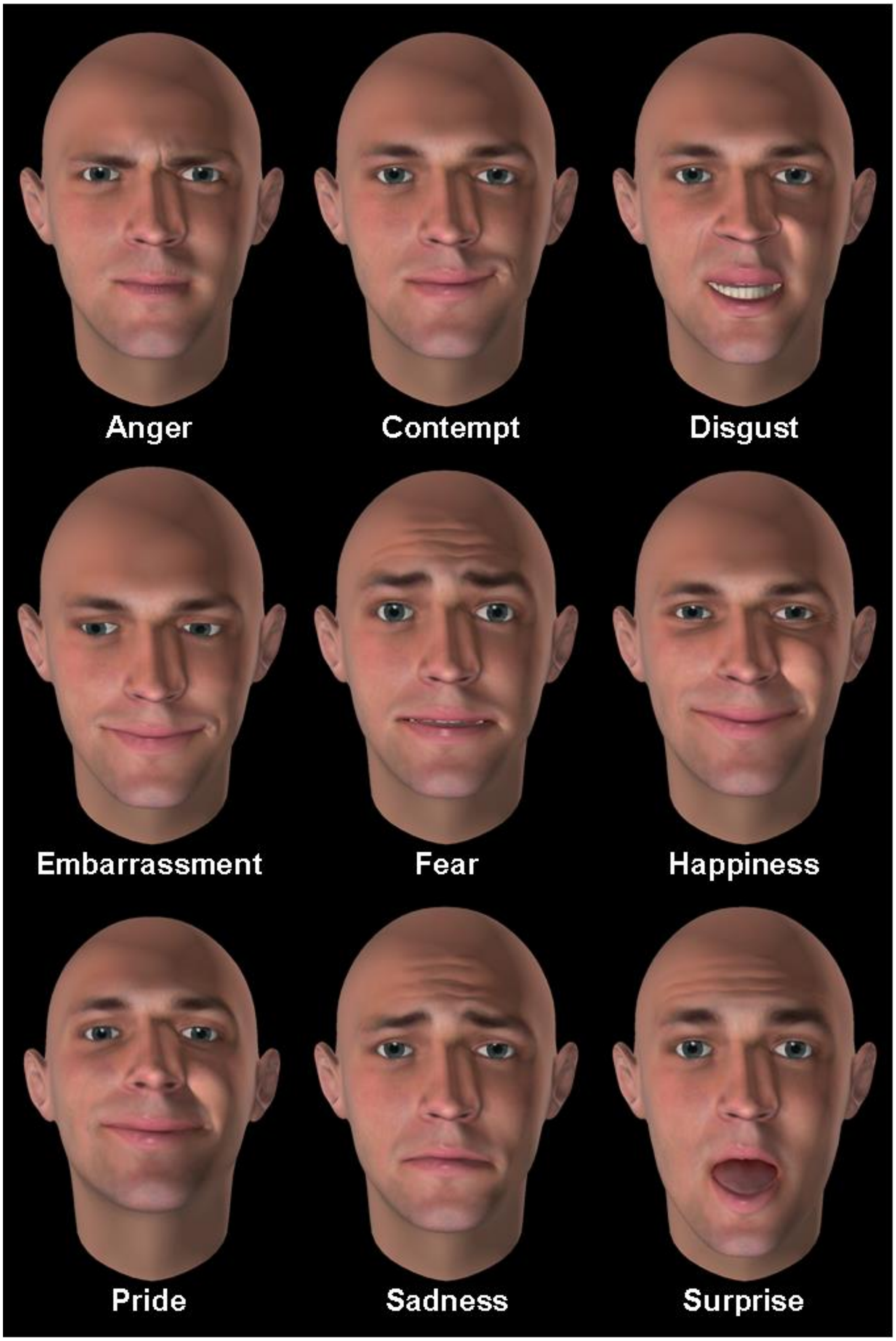

Anger
Contempt
Disgust
Embarrassment
Fear
Happiness
Pride
Sadness
Surprise

Figure 1. Examples of nine emotions as expressed by a photofit FACSGen face at high intensity in the recognition task of Experiment 2. (Emotion labels have been added for illustrative purposes, but were not part of the experimental study.)

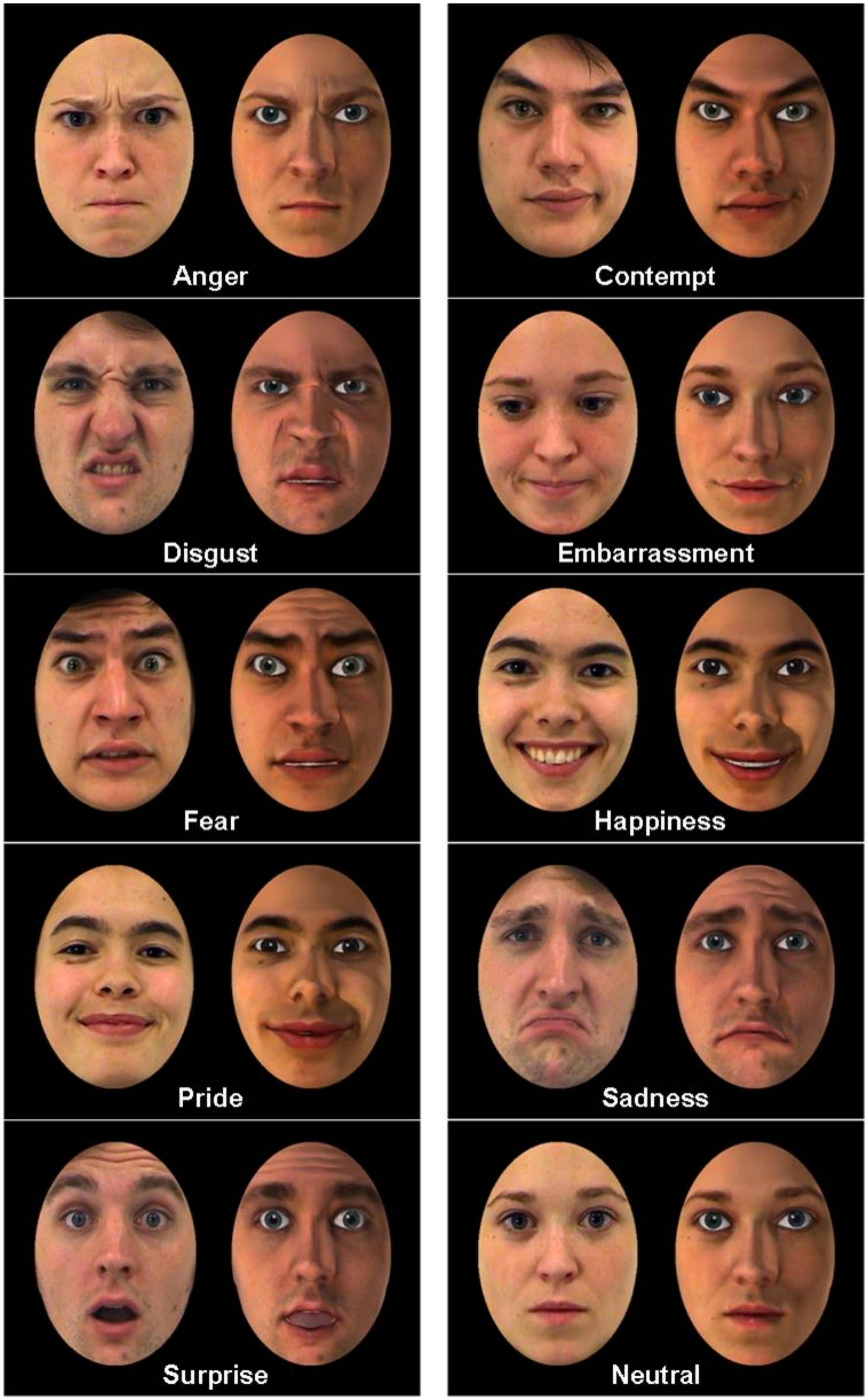

Anger
Contempt
Disgust
Embarrassment
Fear
Happiness
Pride
Sadness
Surprise
Neutral

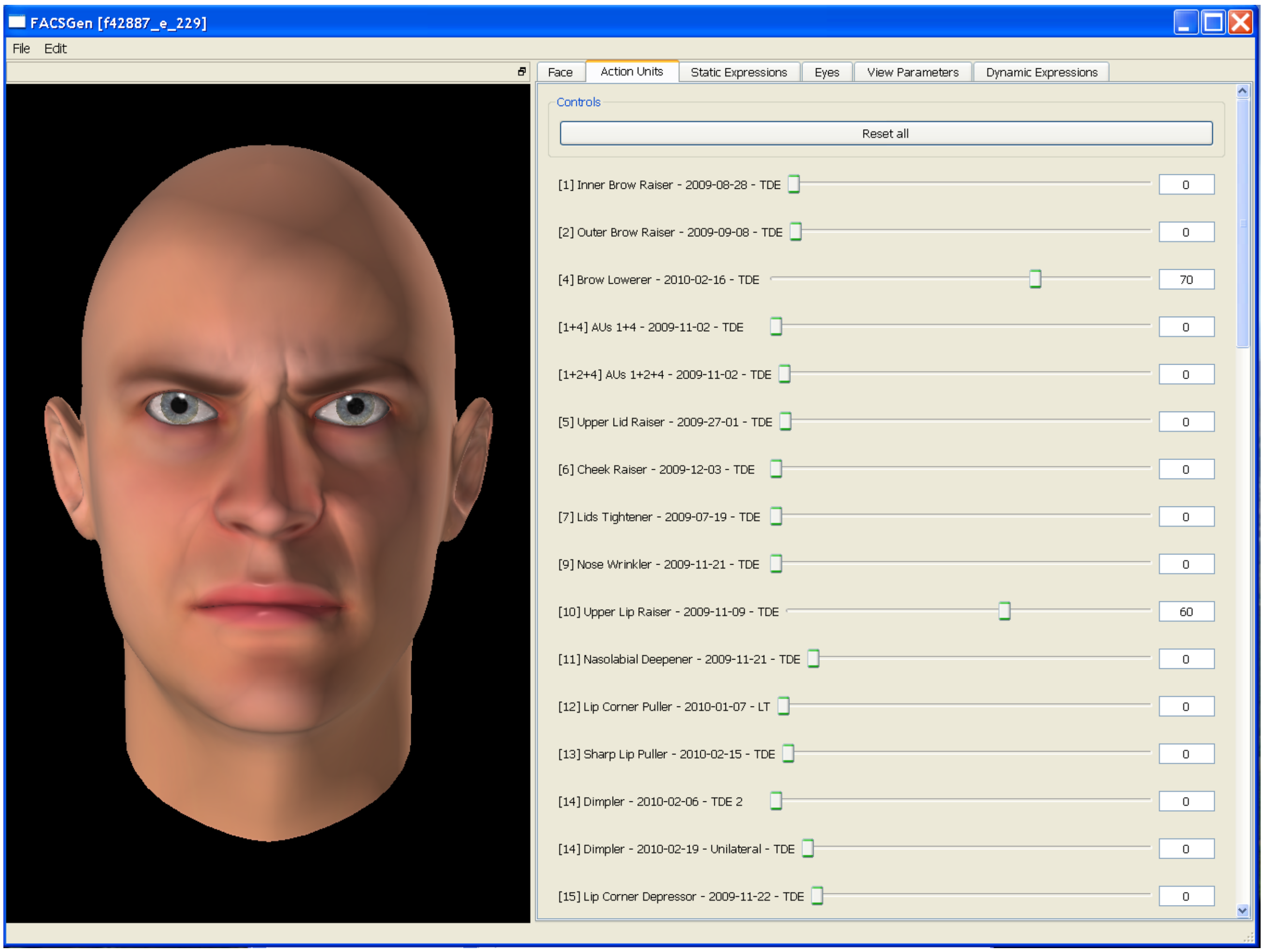
Figure 2. Examples of 10 emotions as displayed by human faces and photofit FACSGen faces in the comparison task of Experiment 2. (Emotion labels have been added for illustrative purposes, but were not part of the experimental study.)

Figure 3. Similarity ratings (1-7) of human and FACSGen expressions for 10 emotions in the comparison task of Experiment 2. Error bars represent standard errors.

Appendix A

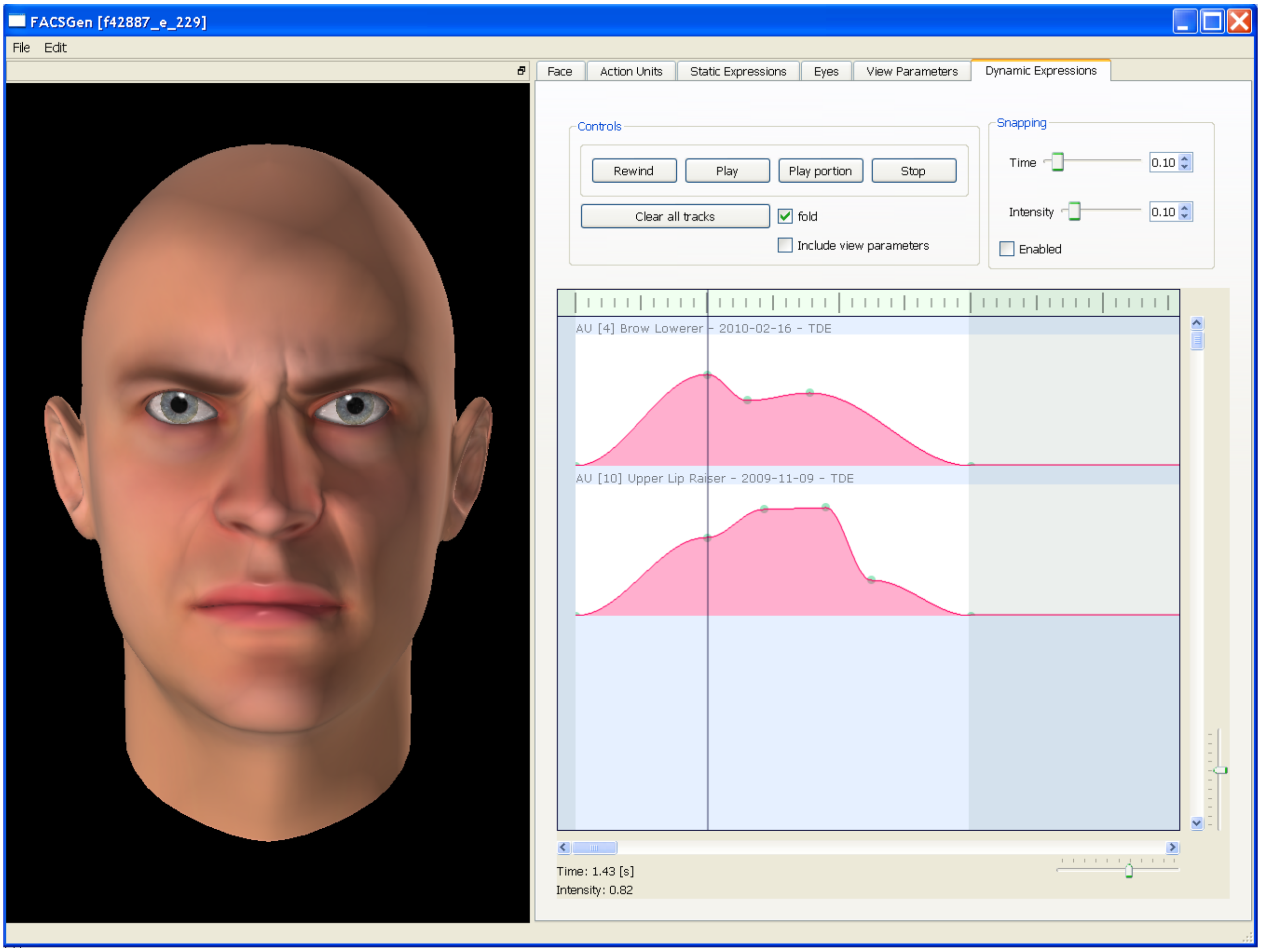
Control Panel in FACSGen 2.0 With Sliders for Each Action Unit (AU) That Can Be Adjusted in Magnitude From 0 to 100%

In the present example, AU4 and AU10 have been activated at 70% and 60% intensity, respectively.



Appendix B

Control Panel in FACSGen 2.0 That Allows the Creation of Dynamic Facial Expressions Over Time Through Nonlinear Manipulation of Activation Curves of Action Units (AUs)



Appendix C

Overview of 35 Single Action Units (AUs) and 54 AU Combinations of the Facial Action Coding System (Ekman, Friesen, & Hager, 2002) as Synthesized by FACSGen 2.0 and Validated in Experiment 1

Single AUs	Name	AU Combinations	
1	Inner Brow Raiser	1+2	22+23+25
2	Outer Brow Raiser	1+4	23+25+26
4	Brow Lowerer	1+2+4	14+17
5	Upper Lid Raiser	1+2+5	14+23
6	Cheek Raiser	4+5	15+17
7	Lid Tightener	5+7	15+23
9	Nose Wrinkler	6+43	17+23
10	Upper Lip Raiser	6+7+12	17+24
11	Nasolabial Furrow Deepener	6+12+15	18+23
12	Lip Corner Puller	6+12+15+17	20+25+26
13	Sharp Lip Puller	6+12+17+23	20+25+27
14	Dimpler (bilateral)	7+12	4+5+7+24 [Anger]
14uni	Dimpler (unilateral)	7+43	10+16+25+26 [Disgust]
15	Lip Corner Depressor	9+17	14+54+62+64 [Embarrassment]
16 (+25)	Lower Lip Depressor	9+16+25	1+2+4+5+20+25+26 [Fear]
17	Chin Raiser	10+14	6+12 [Happiness]
18	Lip Pucker	10+15	12+53/+64 [Pride]

20	Lip Stretcher	10+17	1+4+15 [Sadness]
22 (+25)	Lip Funneler	10+12+25	1+2+5+25+27 [Surprise]
23	Lip Tightener	10+15+17	
24	Lip Presser	10+16+25	
25	Lips Part	10+17+23	
26 (+25)	Jaw Drop	10+20+25	
27 (+25)	Mouth Stretch	10+23+25	
43	Eye Closure	10+12+16+25	
45	Blink	12+15	
46	Wink	12+17	
51	Head Turn Left	12+23	
52	Head Turn Right	12+24	
53	Head Up	12+25+26	
54	Head Down	12+25+27	
61	Eyes Turn Left	12+15+17	
62	Eyes Turn Right	12+16+25	
63	Eyes Up	12+17+23	
64	Eyes Down	20+23+25	

Note. AU25 is automatically scored with AU16 and AU22, which usually part the lips, and with AU26 and AU27, which were implemented as open-mouth actions.