J Nonverbal Behav
DOI 10.1007/s10919-010-0095-9

ORIGINAL PAPER

FACSGen: A Tool to Synthesize Emotional Facial Expressions Through Systematic Manipulation of Facial Action Units

Etienne B. Roesch · Lucas Tamarit · Lionel Reveret · Didier Grandjean · David Sander · Klaus R. Scherer

Abstract To investigate the perception of emotional facial expressions, researchers rely on shared sets of photos or videos, most often generated by actor portrayals. The drawback of such standardized material is a lack of flexibility and controllability, as it does not allow the systematic parametric manipulation of specific features of facial expressions on the one hand, and of more general properties of the facial identity (age, ethnicity, gender) on the other. To remedy this problem, we developed FACSGen: a novel tool that allows the creation of realistic synthetic 3D facial stimuli, both static and dynamic, based on the Facial Action Coding System. FACSGen provides researchers with total control over facial action units, and corresponding informational cues in 3D synthetic faces. We present four studies validating both the software and the general methodology of systematically generating controlled facial expression patterns for stimulus presentation.

Keywords Emotion · Facial expression · Software · Research material · Facial action coding system · FACS

FACSGen is a software developed at the Swiss Centre for Affective Sciences for research purposes. It is only available on a per collaboration basis. More information can be found at http://www.affective-sciences.ch/facsgen. FaceGen Modeller can be purchased from Singular Inversion Inc. Prices and a demonstration version of the software can be found on http://www.facegen.com.

E. B. Roesch · D. Grandjean · D. Sander · K. R. Scherer
Department of Psychology, University of Geneva, Geneva, Switzerland

E. B. Roesch · L. Tamarit · D. Grandjean · D. Sander · K. R. Scherer
Swiss Centre for Affective Sciences, CISA, University of Geneva, 7 rue des Battoirs, 1205 Geneva, Switzerland

L. Reveret
INRIA, Rhône-Alpes, France

E. B. Roesch (✉)
Department of Computing, Imperial College London, 180 Queen's Gate, SW7 2AZ London, UK
e-mail: contact@etienneroes.ch

Published online: 19 November 2010

Much of the research addressing the communication of emotion concerns the perception and interpretation of facial expressions. Typically, participants are shown still pictures or videos of facial expressions, and researchers analyze recognition rates and confusion matrices (e.g., Hess et al. 1997). Alternatively, some researchers, in the field of neuroscience for instance, may be interested in the measurement of the influence of the perception of a given facial expression on a secondary task, like in experiments involving priming (e.g., Ruys and Stapel 2008) or backward masking (e.g., Szczepanowski and Pessoa 2007) of facial expressions.

A growing number of databases are available, containing a large number of facial expressions (e.g., Goeleven et al. 2008; Kanade et al. 2000; Lundqvist et al. 1998; Pantic et al. 2005; or Hirsh et al. 2009 for the facial expression of pain). This material can of course be used as is, but researchers often manipulate it to suit their needs. Specific types of facial expressions can be investigated by applying various methods to create specific experimental stimuli. Image morphing techniques, for instance, allow the creation of dynamic facial expressions by extrapolating a configuration of facial features from a source picture and transfer it to a target picture. Using this technique, a typical stimulus would show a neutral face evolving into one of the basic emotions (e.g., Joorman and Gotlib 2006). However, manipulation of this kind has some limitations related, for example, to the assumption that the actual dynamics of unfolding of emotional expressions can be faithfully represented by a linear function from neutral to emotional expression reflects—an assumption for which there is little evidence (see Scherer and Ellgring 2007). Researchers can manipulate this tailor-made material in certain ways, for example, by specifying the speed of the unfolding, or creating ambiguous stimuli midway between two emotions.

Whereas shared and standardized stimulus sets facilitate the comparison of results across studies, often researchers are rather limited in their ability to manipulate stimulus features, and to ascertain appropriate experimental control. For instance, only very few databases of actor-posed facial expressions contain facial configurations that have been controlled in terms of precise muscle movements, specifying the informational cues available in the face. Generally, actors are only provided with verbal labels of emotions—often only basic emotions—and instructed to pose the respective expressions corresponding to their personal interpretation of those emotions. A researcher seeking to manipulate particular facial features (e.g., the amount of eye opening for example), or interested in studying less orthodox facial expressions (e.g., the expression of pain), is thus left with the difficult task of creating a dedicated database.

An alternative approach is to use computer-generated facial expressions. In recent years, facial animation attracted a lot of attention in the computer graphics community (Parke and Waters 1996). A number of successful solutions have been proposed (e.g., Bickel et al. 2007; Blanz and Vetter 1999; Cosker et al. 2008; Ma et al. 2008; Zhang et al. 2004), making use of motion capture techniques to record facial expressions portrayed by actors, and developing algorithms to render high quality animations reproducing the facial expressions recorded. Once synthesized, these facial expressions could, theoretically, be manipulated to create ad hoc facial expressions, in much the same vein as modern animated Hollywood movies. However, despite the high quality graphics that can be produced, the technical investment is such that it is very unlikely to appeal to researchers in psychology or neuroscience who seek to produce material to investigate facial expressions.

In this article, we describe FACSGen, a novel tool to create experimentally controlled facial expression patterns. FACSGen takes advantage of the flexibility of the Facial Action Coding System (FACS; Ekman and Friesen 1976; Ekman et al. 2002) to represent facial expressions while fulfilling the requirements for reproducibility and comparison of material across studies. FACSGen allows the creation of realistic facial expressions by parametrically manipulating action units (AUs) portrayed on an infinite number of synthesized 3D faces, created with FaceGen Modeller ([Software] 2009)—see for instance Cristinzio et al. (2010), N'Diaye et al. (2009), Roesch et al. (2009, 2010a, b) for examples of AUs manipulation using FACSGen in various experimental settings. The FACS defines the common AUs that facial muscles can produce, thus allowing the formal description of the constituents of any facial expression. It contains 58 AUs, out of which 20 are commonly used to describe most facial expressions of emotions. FACSGen dynamically combines individual AUs to generate a virtually infinite variety of dynamic facial expressions and allow the creation of an unlimited number of facial expressions, static or dynamic, that can be modeled on a potentially infinite number of facial identities. Thus, FACSGen promises to become a key tool in the investigation of the perception of facial expressions in general, and the inferences from emotional facial expressions in particular.

We begin by introducing FaceGen Modeller, a commercial tool we use to create and handle realistic 3D facial stimuli. We then describe FACSGen, the tool we developed, which allows the parametric manipulation of facial AUs, as an add-on to FaceGen Modeller. Next, we present four studies that we conducted to validate the software as well as the methodology of using the synthetic stimuli created with this tool. We conclude by discussing the potential of FACSGen as compared to other software currently available.

FACSGen: The Parameterization of Facial Expressions

FACSGen is a software we developed to manipulate the expression of synthesized 3D faces on the basis of the FACS. It is used in conjunction with FaceGen Modeller (2009), a commercial software that allows the creation and manipulation of an infinite number of realistic synthesized faces (Corneille et al. 2007; Freeman and Ambady 2009; Moradi et al. 2005; Oosterhof and Todorov 2008; Schulte-Rüther et al. 2007; Shimojo et al. 2003; Todorov et al. 2008). A representation of the information flow in FACSGen is shown in Fig. 1. As can be seen, FACSGen integrates components of FaceGen Modeller, allowing the precise and coherent control of the same 3D objects, and providing FACSGen with many of the features available in FaceGen Modeller. Basically, the user provides FACSGen with a representation of the dynamics of activation over time for each AU (the curves in the figure), and FACSGen produces the corresponding facial expression unfolding over time, either in the form of a series of still pictures (png files) or a movie (avi files). This facial expression can be mapped onto any face created using FaceGen Modeller.

FaceGen Modeller

FaceGen Modeller is a commercial tool that was originally designed for the creation of realistic 3D faces in video games (Singular Inversions Inc. 2009). It is based on a database of thousands of human faces whose shape and texture have been digitized into 3D objects. Representing these faces into factorial space allows the extrapolation of new, unique, faces on the basis of a number of continua. Faces created with FaceGen Modeller vary in gender, age, and ethnicity, and can be manipulated in ways very similar to a sculptor to create very realistic faces. All faces from all constitutions (e.g., chubby, skinny) and all shapes (e.g., sharp-edged, oval) could, in theory, be reproduced. The user interacts with the software through an intuitive graphical user interface, without requiring special training. FaceGen Modeller also provides the user with the ability to create their own 3D mesh (i.e., 3D

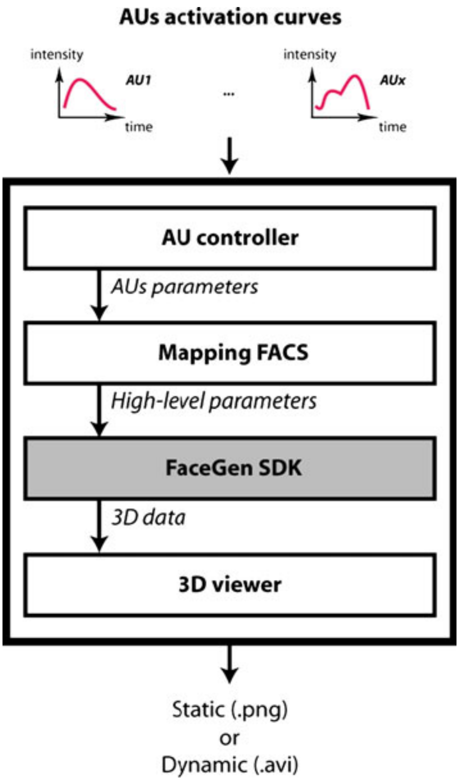

Fig. 1 Architecture and information flow in FACSGen. A graphical user interface allows the user to describe the dynamics of activation over time for each AU (curves at the top of the figure). Different layers process this information, mapping it to action units, before applying the corresponding facial expression on 3D faces created with FaceGen Modeller. FACSGen manipulate FaceGen faces through the FaceGen software development kit released by Singular Inversion Inc

topology and detailed texture) from close-up photographs of a person. Digitized faces can be altered and imported into FACSGen just like any other 3D faces created using FaceGen Modeller (Fig. 2). By default, heads created with FaceGen Modeller are bald, but additional 3D objects (e.g., hair, facial hair, or miscellaneous accessories) can be added if needed. FaceGen Modeller is primarily dedicated to the creation of 3D facial morphology. Given a specific morphology, FaceGen Modeller allows limited control over the manipulation of some basic features of facial expression (e.g., gaze, head direction, or morphological changes due to phonology) and offers a small number of full-blown, non FACS-based emotional expressions. In our first study, because researchers already using FaceGen Modeller may want to use these built-in expressions, we asked FACS coders to code these expressions (denoted "FG expressions" in the article) as well as expressions produced using FACSGen.

FACSGen

FACSGen is a software that can import any face exported from FaceGen Modeller (i.e., created from scratch or from close-up photographs; see Fig. 2). It interfaces with FaceGen Modeller through a C++ SDK library released by Singular Inversion Inc. that allows the manipulation of the modeled 3D objects. The SDK provides access to 150 high-level, morphological parameters manipulating different aspects of the topology of the face (see Fig. 1). In some cases, FACS coders are required to base their judgment on both the movements performed by the muscles of the face and the co-occurrence of particular

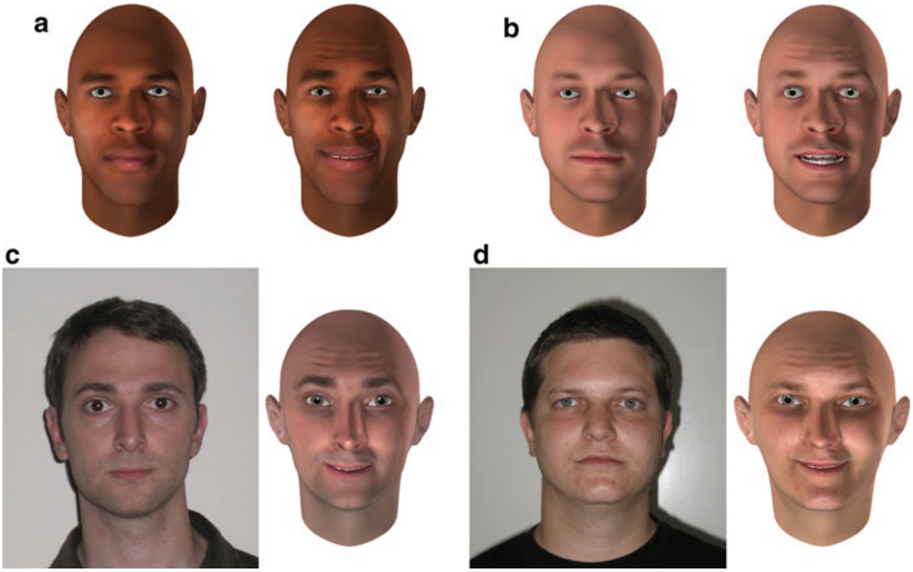

Fig. 2 Examples of faces manipulated with FACSGen. *Panel a* Random African face: neutral expression and portraying AU 1 + 2 + 6 + 12 + 25. *Panel b* Random Caucasian face: AU 1 + 2 + 5 + 25 + 26. *Panel c* and *d* Synthesized faces digitized from close-up pictures: neutral photograph and portraying the exact same facial expression (AU 1 + 2 + 6 + 12 + 25)

features, like the wrinkling of the skin and changes in its pigmentation. In the presence of AU 12 "smile", for instance, FACS coders will code the activation of AU 6 "cheek raise" if so-called crow's feet wrinkles appear in the outer corner of the eyes. Situations of this type not only involve changes in the morphology of the face but also in the visual aspect of the skin. As FaceGen Modeller itself does not support the manipulation of such features, we complemented FaceGen parameters with our own set of dedicated parameters and augmented graphical representations.

In FACSGen, a graphical user interface allows both the linear manipulation of AUs (Fig. 3) to edit a static face, and the non-linear manipulation of activation curves (Fig. 4), which allow the representation of complex dynamic changes over time. The visual output consists of a sequence of frames depicting the unfolding facial expression by mapping the intensity for each AU and for each point in time. These frames can then be used individually as static displays of an evolving facial expression, or sequentially composed into a movie clip. The scalar values of the activation curves can be exported in text files for offline analyses, and imported back again in FACSGen to generate the exact same facial expressions on different faces. This feature responds to the need for the reproducibility of experimental setups across studies.

Validation Studies

The general methodology consists in creating a number of faces using FaceGen Modeller to define the base morphology, importing and manipulating them in FACSGen to create controlled facial expressions to be used as experimental material for the systematic study

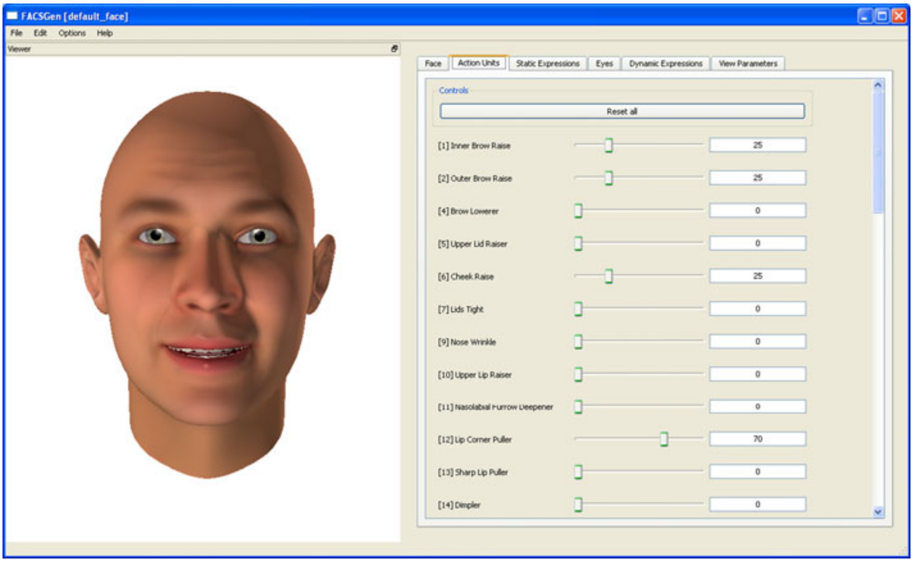

Fig. 3 Screenshot of FACSGen. This panel allows the user to create dynamic facial expressions following linear trends similar to morphing techniques. The resulting expression can be exported as a movie clip, and each step of the unfolding can be exported as a still picture

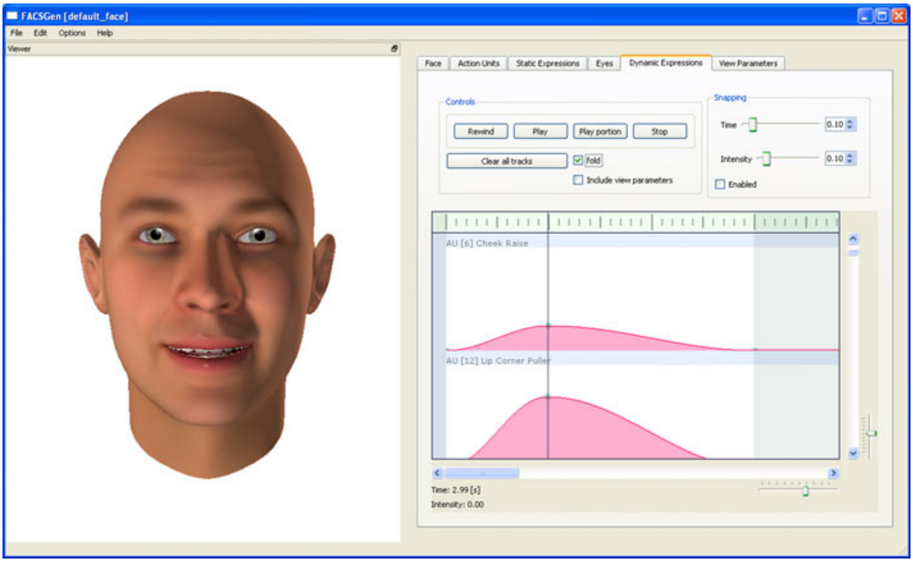

Fig. 4 Screenshot of FACSGen. This panel allows the user to create dynamic facial expressions following non-linear trends. The resulting expression can be exported as a movie clip, and each step of the unfolding can be exported as a still picture

of inferences made from AU static and/or dynamic configurations. The creation of ecologically valid stimuli requires that (a) the AUs manipulated in FACSGen correspond to what is described in the FACS, (b) synthesized 3D identities produced in FaceGen

Modeller are reliably perceived by lay participants (e.g., in terms of gender, believability, and intrinsic emotionality, related for example to attractiveness or trust), and (c) that the manipulation of AUs on these identities produce facial expressions that are reliably recognized by lay participants. Here, we consider these as the central criteria for a validation of the tool.

The validation procedure we conducted consisted of three parts. In study 1, we verified that our operational definitions of the AUs, and their implementation on a 3D model, correspond to convergent coding by trained FACS coders. For this study, we created a number of video clips depicting combinations of AUs that were submitted to certified FACS coders for coding. In study 2, and studies 3a and 3b, we verified the applicability of the FACSGen modeling procedure to frequently encountered experimental settings using static pictures. Specifically, in study 2, we investigated how lay participants perceived expressive faces, with neutral faces, using 3D synthesis with FaceGen Modeller, to examine the quality of the facial identities In studies 3a and 3b, we manipulated a certain number of these faces using FACSGen, applying a selection of the controlled emotional facial expressions validated in the first study, and asked participants to rate the underlying emotion. In this part of the research, we used two different versions of these emotional faces: color stimuli (study 3a) and processed grayscale stimuli (study 3b). This comparison was made because a growing number of researchers, especially in psychophysics and neuroscience, seek to control for the low-level features of visual material, including facial expression, by manipulating it in a number of ways. For instance, often color pictures of facial expressions are converted into grayscale, and the gray level intensity of all pixels is normalized to ensure that the recognition of emotions is not biased by the general luminance and contrast of a particular experimental condition (e.g., Pourtois et al. 2004, see also Delplanque et al. 2007 for related issues involving spatial frequencies). Because such widely used techniques may alter the general perception of facial expressions—which might pose particularly serious problems for synthesized 3D faces—we had both color and processed grayscale pictures separately evaluated by lay judges.

Study 1: Validation of the Action Units Manipulated by FACSGen

The facial action coding system (FACS, Ekman and Friesen 1976; Ekman et al. 2002) is the most popular facial coding scheme to date. It defines common action units that facial muscles can operate, and thus allows the formal and exhaustive description of the constituents of any facial expression, static or dynamic. Comparing the results produced with FACSGen to the AUs defined in the FACS manual is therefore a critical step in the evaluation of our software.

Procedure

Four certified FACS coders were provided with 2 s clips (50 frames) depicting the unfolding of either a single AU or a combination of several AUs, evolving from no activation to 100% of activation to produce realistic full-blown expressions. Clips were numbered, and each FACS coder was assigned to a randomized presentation order. For each clip, FACS coders were instructed to mark if the AUs were present (noted 1) or absent (noted 0), yielding an activation profile of AUs. FACS coders were not asked to rate the intensity of activation for each AU, as judgments of this parameter show poor inter-rater agreement (Sayette et al. 2004).

Each AU was portrayed on 6 different Caucasian identities (3 females). Portrayals unfolded from a neutral expression to one out of 16 single AUs, or one out of 17 combinations of AUs as described by the Emotional FACS (EMFACS; Ekman et al. 1994). In addition, because emotional facial expressions created with FaceGen Modeller have already been used in research, and are likely to be integrated in setups involving FACSGen, we created and evaluated clips portraying FaceGen built-in facial expressions (denoted "FG"). All clips portrayed one face in full color against a black background, frontal view, and facing the observer.

Results and Discussion

Table 1 shows that both the clips portraying a single AU, and the clips portraying a combination of AUs were reliably recognized and coded by FACS coders. Cronbach α were computed for each FACSGen manipulation, using coded AUs as items, and the 24 profiles (4 coders $\times$ 6 identities) of AUs as variables. We did not manipulate and evaluate AUs that describe head movements. Consequently, all faces were facing the observer, which may have had an effect on the coding of certain AUs. For instance, FACS coders rarely coded AU 19 "tongue show", but mostly coded AU 25 + 26 describing the opening of the mouth, even though the tongue was visible. On the whole, we conclude that certified FACS coders reliably recognize the synthetically produced AUs and AU configurations.

Study 2: Establishing the Ecological Validity of Facial Identities Produced Using FaceGen Modeller

FaceGen Modeller can create an infinite number of synthesized 3D identities, either from scratch or from digitized close-up photographs (see Fig. 2). To use this computer-generated material to systematically study the interaction between facial expression and facial identity, we need to ensure that lay participants reliably perceive the synthesized identities. In an optimal situation, faces produced with FaceGen Modeller should be unambiguously recognized as male or female (note that some researchers may need androgynous faces, in which case the procedure would have to be adapted accordingly). They should also be most believable (i.e., looking as natural as possible given the limitations of 3D synthesis) and as emotionally neutral as possible. Study 2 addressed this issue, and allowed selecting a pool of rated identities from which we selected the faces used in studies 3a and 3b.

Procedure

We created 180 faces using FaceGen Modeller: Faces were Caucasian male or female, of an estimated age between 20 and 45 years old. The faces were created with the aim of being as believable as possible, and as emotionally neutral as possible. Color pictures of the faces were then presented in random orders to 44 students (35 females, mean age 23.7 years) from the University of Geneva. Participants were gathered in classrooms, and used a web-based interface to report their judgments. Participants were instructed to rate the 180 faces on three continuous dimensions: gender—"Is this the face of a male, an androgynous person, a female?" (anchored "*Male*", "*Female*"), believability—"Is this face natural? Could you encounter it in the street?" (anchored "*Synthetic*", "*Believable/Realistic/Natural*"), and intrinsic emotionality—"Does this face seem to show a positive, neutral, or negative emotion?" (anchored "*Positive*", "*Neutral*", "*Negative*").

Table 1 Inter-rater agreement and results of coding by FACS coders

FACSGen manipulation	Name	α	AUs coded
AU 1	Inner brow raiser	.990	AU 1
AU 2	Outer brow raiser	.989	AU 2
AU 4	Brow lowerer	.998	AU 4
AU 6	Cheek raiser	.975	AU 6
AU 7	Lids tight	.982	AU 7
AU 12	Lip corner puller	.987	AU 12 (+ 6)
AU 17	Chin raiser	.977	AU 17 (+ 5 + 24)
AU 19 (+ 25 + 26)	Tongue show	.988	AU 25 + 26 (+ 10)
AU 20	Lip stretch	.903	AU 20 (+ 12 + 24)
AU 22	Lip funneler	.992	AU 22 + 25
AU 23	Lip tightener	.933	AU 23 (+ 5 + 24)
AU 25 + 26	Jaw drop	.998	AU 25 + 26
AU 61	Eyes left	.989	AU 61
AU 62	Eyes right	.991	AU 62
AU 63	Eyes up	1.	AU 63
AU 64	Eyes down	.998	AU 64
AU 1 + 2		.994	AU 1 + 2 (+ 5)
AU 1 + 2 + 5	Surprise	.982	AU 1 + 2 + 5
AU 1 + 2 + 5 + 25 + 26	Fear	.994	AU 1 + 2 + 5 + 25 + 26
AU 4 + 7	Anger	.993	AU 4 + 7
AU 4 + 7 + 23	Anger	.990	AU 4 + 7 + 23
AU 4 + 17 + 23	Anger	.978	AU 4 + 5 + 23
AU 5 + 25 + 26	Surprise	.997	AU 5 + 25 + 26
AU 12 + 25	Happiness	.987	AU 12 + 25 (+ 6)
AU 22 + 25 + 26	Neutral (mouth open)	.970	AU 22 + 25 + 26
FG: Anger		.971	AU 9 + 16 + 25
FG: Anger + AU 25	Anger	.968	AU 9 + 25 + 26 (+ 16)
FG: Disgust		.988	AU 9 + 25 (+ 4 + 10)
FG: Fear		.978	AU 1 + 25 + 26 (+ 4 + 7 + 10)
FG: Sadness		.986	AU 4 + 7 (+ 24)
FG: Surprise		.985	AU 1 + 2 + 5 + 25
FG: Fear + AU 1 + 2 + 5 + 25 + 26	Fear	.994	AU 1 + 2 + 5 + 25 + 26
FG: Hap + AU 1 + 2 + 6 + 12 + 25	Happiness	.992	AU 1 + 2 + 6 + 12 + 25

Parentheses show AUs that were proposed by some but not all FACS coders

Results and Discussion

Cronbach α were computed for each of the three dimensions, using participants' ratings as columns (items) and the 180 pictures as rows (cases). Single measures intra-class correlation coefficients are indicated in parentheses. Results showed that the faces were reliably rated on the three dimensions: α for gender = 1.00 (*ICC* = .851); for believability = 0.94 (*ICC* = .265); and for intrinsic emotionality = 0.96 (*ICC* = .376). To determine whether participants managed to discriminate the faces on the three dimensions, a *t*-test was performed on the ratings obtained for each of the three dimensions, comparing the first and last quartiles of the respective ordered sample. Results showed that the faces could be discriminated on each of the dimensions. Male faces yielded ratings significantly closer to the "Male" anchor ($M = 2.79$, $SD = 4.11$) than did female faces ($M = 81.99$, $SD = 17.76$), $t(2141) = -191$, $p < .001$. Results also showed that the first quartile of the sample was significantly less believable ($M = 29.67$, $SD = 27.94$) than the last quartile of the sample ($M = 68.63$, $SD = 28.87$), $t(3820) = -42.3$, $p < .001$. Finally, results showed that the first quartile of the sample was perceived more negatively ($M = 35.72$, $SD = 16.11$) than the last quartile of the sample ($M = 62.51$, $SD = 15.08$), $t(3790) = -53.0$; $p < .001$.

Overall, these results show that lay participants reliably perceive the gender of FaceGen faces, and can reliably attribute ratings of believability, and intrinsic emotionality to such faces. Because FaceGen allows the creation of very different kinds of faces (from very realistic to more caricature-like), it is very important to be able to assess and control these dimensions.

Study 3a: Validation of Emotion Inferences Drawn from FACSGen Facial Expressions (Color Version)

Procedure

Out of the 180 faces created for Study 2, we selected 77 faces (40 females), for being the most unambiguously gender-typed, the most believable, and the most emotionally neutral faces. We then manipulated the faces using FACSGen, to depict the combinations "Anger: AU 9 + 16 + 25", "Fear: AU 1 + 2 + 5 + 25 + 26", "Happiness: AU 1 + 2 + 6 + 12 + 25", and "Neutral: AU 22 + 25 + 26" (as described in Study 1, and validated by FACS coders). These AU combinations do not fully concur with some of the complete prototypical facial expressions described in the literature (although there is much discrepancy in these descriptions and complete prototypical configurations are very rarely found, see Scherer and Ellgring 2007). However, these combinations are very likely to occur in real life situations and frequently occur in actor portrayals of emotions (see Scherer and Ellgring). In consequence, we assumed that they can be recognized by lay participants with reasonable agreement.

The procedure used was similar to Study 2. Twenty students (14 females, mean age 22.1 years) rated color pictures of the 77 faces, each of which portrayed the four facial expressions. Participants were instructed to rate the extent to which the following emotions could be perceived in the facial expressions: anger, disgust, fear, happiness, sadness, and surprise. Blank fields allowed them to propose other emotions. A scale from 0 (*anchored "not at all"*) to 100 (*anchored "enormously"*) was provided. They also had to rate the overall intensity of the facial expression. A scale from 0 (*anchored "not intense"*) to 100 (*anchored "very intense"*) was provided.

Results and Discussion

Cronbach α were computed, using participants' ratings as columns (items), and pictures as rows (cases). Single measures intra-class correlation coefficients are indicated in parentheses. Results showed that the faces were reliably rated on the 7 scales (6 emotions, and intensity): α for anger = 0.98 ($ICC = .63$); for disgust = 0.92 ($ICC = .30$); for happiness = 0.97 ($ICC = .55$); for fear = 0.98 ($ICC = .62$); for surprise = 0.82 ($ICC = .19$);

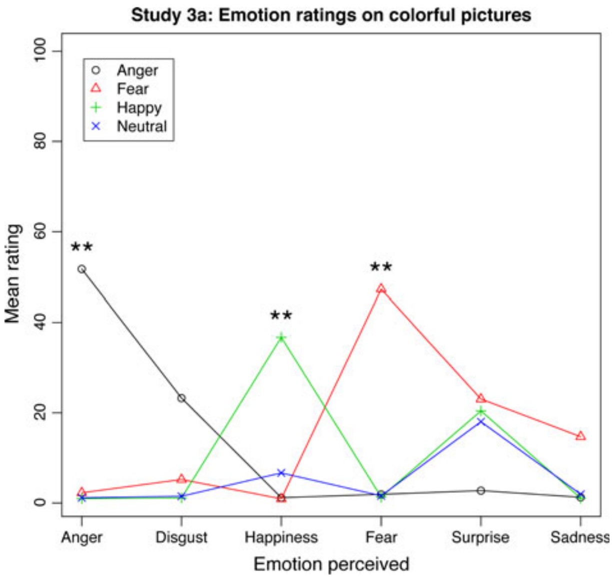

Fig. 5 Results of study 3a. Emotion ratings on *color pictures*. Faces were depicting the FACSGen combinations: "Anger: AU 9 + 16 + 25", "Fear: AU 1 + 2 + 5 + 25 + 26", "Happiness: AU 1 + 2 + 6 + 12 + 25", and "AU 22 + 25 + 26" (as described in Study 1, and validated by FACS coders)

for sadness = 0.87 (*ICC* = .24); and for intensity = 0.94 (*ICC* = .39). To determine whether participants can discriminate the emotions portrayed by FACSGen faces portrayed on color pictures, repeated measures analysis of variance (ANOVAs) was performed for each of the three portrayed emotions anger, fear, happiness and neutral. The dependent variables were participants' ratings on the 7 scales. In all four cases, there was a significant main effect of emotion (anger: $F(5,380) = 1337.0$, $p < .001$; fear: $F(5,380) = 784.7$, $p < .001$; happiness: $F(5,380) = 1010.0$, $p < .001$; neutral: $F(5,380) = 276.9$, $p < .001$). Contrast analyses were performed by assigning the target emotion as the reference category. There was a significant effect for targets' emotions (*p*s < .001), indicating that participants reliably recognized the target emotion portrayed by the faces (see Fig. 5). The mean intensity ratings for anger, fear, happiness and neutral were 60 (SD = 5.52), 55.6 (SD = 5.82), 37.9 (SD = 4.78) and 20.8 (SD = 5.52) respectively.

We conclude that lay participants can recognize, with satisfactory accuracy and a very high degree of agreement, the emotions represented by specific AU configurations in FACSGen produced facial expressions, built after consensual descriptions found in the literature, as well as some empirical findings (Scherer and Ellgring 2007).

Study 3b: Validation of Emotion Inferences Drawn from FACSGen Facial Expressions (Grayscale Version)

Procedure

Whereas most psychologists investigating the perception and interpretation of facial expressions do not need to alter the stimulus sets available to the research community, a growing number of researchers in adjacent disciplines—especially in neuroscience and psychophysics—seek to disentangle the higher-level effects of the facial expressions from the lower-level effects of several dimensions intrinsic to the visual material depicting the facial expressions; e.g., the saliency and contrast (e.g., Pourtois et al. 2004) or the spatial frequencies (e.g., Delplanque et al. 2007). To do so, they process this material in ways that may affect its recognition (Fiser et al. 2003). Because this situation may be even worse for synthesized material, like faces produced with FaceGen Modeller, we replicated the evaluation of Study 3a with a sub-selection of the 77 emotional facial expressions that were being used in a separate experiment, and which we processed in the same way as Pourtois et al.

Study 3b was part of larger-scale experiment investigating how the perception of the emotions of fear and happiness modulate attention (the results of this work are reported in Roesch et al. 2009, 2010a, b). Ten faces (5 females) were selected from the panel of 77 faces created for Study 3a. The faces depicted combinations of AUs representing fear, happiness, and neutral. Because it was irrelevant for the above-mentioned experiments, anger was not included. Each face was converted to grayscale and analyzed in Matlab to extract the mean pixel luminance and contrast range. Statistical analyses confirmed that the three emotional conditions did not differ for luminance and contrast (Pourtois et al. 2004). Upon completion of the attention-emotion experiment, participants rated the faces they had seen during the experiment in a procedure similar to Study 3a: 37 students (29 females, mean age 22.9 years) rated the grayscale pictures of the 10 faces (5 females), each of which portrayed the three facial expressions.

Results and Discussion

Cronbach α were computed, using participants' ratings as columns (items), and pictures as rows (cases). Single measures intra-class correlation coefficients are indicated in parentheses. Results showed that the faces were reliably rated on the 7 dimensions: α for anger = 0.82 ($ICC = .12$); for disgust = 0.91 ($ICC = .23$); for happiness = .99 ($ICC = .79$); for fear = 0.99 ($ICC = .75$); for surprise = 0.82 ($ICC = .1$); for sadness = 0.93 ($ICC = .24$); and for intensity = 0.97 ($ICC = .43$). Means on the ratings showed that participants perceived the intended emotions in the facial expressions produced (Fig. 6). To determine whether participants can discriminate the emotions portrayed in the grayscale version of the FACSGen faces, repeated measures analysis of variance (ANOVAs) was performed for each of the three portrayed emotions fear, happiness, and neutral. The dependent variables were participants' ratings on the 7 scales. In all four cases, there was a significant main effect of emotion (fear: $F(5,1845) = 399.3$, $p < .001$; happiness: $F(5,1845) = 1048$, $p < .001$; neutral: $F(5,1845) = 157$, $p < .001$). Contrast analyses were performed by assigning the target emotion as the reference category. There was a significant effect for targets' emotions ($p < .001$), indicating that participants reliably recognized the target emotion portrayed by the faces. The mean intensity ratings for fear, happiness, and neutral were 76.99 (SD = 16.78), 60.01 (SD = 20.40) and 36.66 (SD = 23.15), respectively.

We conclude that, from processed grayscale FACSGen stimuli, lay participants can recognize the emotions represented by specific AU configurations with satisfactory accuracy and with a very high degree of agreement.

Summary and Discussion

Researchers interested in facial expressions of emotions often rely on shared sets of stimuli. This material contains static pictures, or videos of facial expressions portrayed by

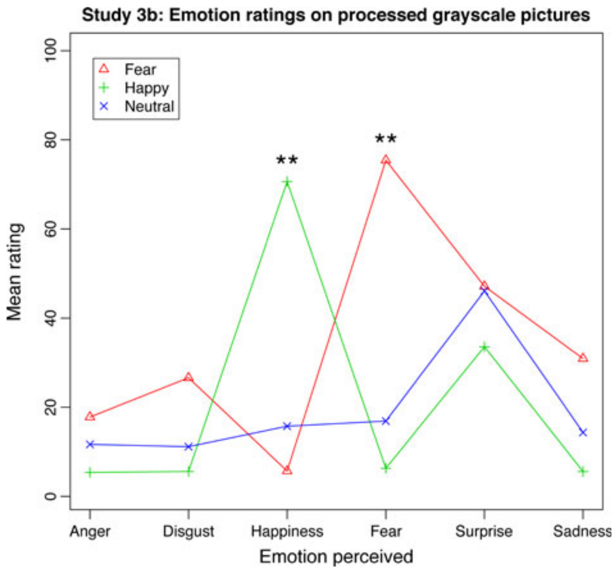

Fig. 6 Results of study 3b. Emotion ratings on processed *grayscale pictures*. Faces were depicting the FACSGen combinations: "Fear: AU 1 + 2 + 5 + 25 + 26", "Happiness: AU 1 + 2 + 6 + 12 + 25", and "AU 22 + 25 + 26" (as described in Study 1, and validated by FACS coders)

actors or recorded from live occurrence. Although using this shared material has proven useful to compare results across studies, its lack of flexibility often does not allow the necessary degree of experimental control. To avoid creating their own dedicated stimulus sets, researchers have attempted to use techniques such as morphing existing photographs. Only fairly recently have researchers in psychology and neuroscience discovered the potential of synthetic stimuli (e.g., Cristinzio et al. 2010; Ellison and Massaro 1997; Freeman and Ambady 2009; Gaag et al. 2007; Hirsh et al. 2009; Moser et al. 2007; N'Diaye et al. 2009; Parr et al. 2008; Pelphrey et al. 2004; Roesch et al. 2009, 2010a; Sander et al. 2007; Wehrle et al. 2000). In response to the manifest need to develop means of systematically manipulating facial expressions to allow a optimal degree of experimental stimulus control, we developed the tool described in this article, FACSGen.

To validate FACSGen, we investigated the perception of single AUs, combinations of AUs, and complex full-blown facial expressions of emotion that used 16 AUs in total. In four studies, we submitted this material to both certified FACS coders, and lay participants. Our results showed that (a) the AUs manipulated in FACSGen correspond to the FACS specifications and are reliably recognized by certified FACS coders, (b) the synthesized identities produced by FaceGen Modeller are perceived by lay participants as reasonably believable, and thus can be used in place of naturalistic portrayals, and (c) the manipulation of FaceGen faces in FACSGen produces facial expressions of emotions that are reliably recognized by lay participants.

FACSGen can be compared to other software like Poser 7 (2007), Greta (Malatesta et al. 2006; Pasquariello and Pelachaud 2001), realEmotion (Grammer et al. in preparation), or the Virtual Actor Project (Hezle et al. 2004). These solutions differ widely with respect to user friendliness: Researchers often have at their disposal tools that are either too difficult to use in a research context (but produce Hollywood-class material), or very easy to use but produce relatively low quality, often caricature-like material. In contrast, FACSGen has been designed for researchers with the aim to strike a balance between usability and believable realism: on the one hand, FACSGen does not require users to acquire new technical knowledge, compared to the other solutions; on the other, it produces high quality, ecologically valid research material.

Contrasting with other current methods for producing synthetic tailored facial expressions to FACSGen, we identify a number of additional benefits associated with our approach. First, FaceGen Modeller allows the creation of a virtually infinite number of realistic identities, of any gender, age, or ethnicity. For example, the software makes it possible to create androgynous faces, and/or mixtures of ethnic backgrounds. Second, FaceGen Modeller also provides the ability to create 3D meshes from close-up photographs of any person. Digitized faces can then be altered, and used in FACSGen just like any other 3D faces created using FaceGen. Third, the output of FACSGen consists of a series of frames, depicting the unfolding facial expression mapping the activation of AUs on the geometry of the face. The frames can then be used either as still portrayals, or converted into movie clips, overcoming the limitations of morphing techniques. Finally, the activation curves describing the unfolding intensity of each AU in FACSGen can be exported in separate files. These files can be imported back into FACSGen, and applied to a different set of faces or facial expressions, thus allowing comparable material to be used in experimental studies. These options are not available in any of the other software currently available.

Whereas FACSGen provides researchers with new avenues for creating ad hoc material, any synthesized material admittedly poses limitations. First, the information conveyed by facial expressions cannot be reduced to a combination of topological changes in the face. Other channels of information include, for instance, changes in the color, the texture, and the elasticity of the skin; all of which is also subject to great inter-individual differences. These are problematic issues for any synthesis system. FACSGen does, however, take some of these aspects into account in the form of dedicated parameters to create realistic wrinkles, and we are developing more parameters to achieve the best results. Second, through FaceGen Modeller, users of FACSGen can animate 3D models digitized from close-up photographs of a person (see Fig. 2).

To conclude, we presented FACSGen, a novel tool that allows researchers on facial expressions in general, and on facial expressions of emotions, in particular to manipulate informational cues portrayed in facial expressions. It generates synthetic, yet realistic, 3D faces used to produce either static or dynamic material. It also offers a very handy way of representing the unfolding dynamics of the constituents of facial expressions, which allows (1) the portrayal of complex dynamic facial expressions, and (2) the comparison of the material produced between studies. We believe that this new research technology allows researchers to produce appropriate stimulus material for targeted studies to examine specific hypotheses on the AU components of facial expression and their cue value for emotion recognition, constituting a precious tool for critical comparisons between competing theories as well as theory development.

Acknowledgments This work was partially supported by the following sources: HUMAINE, 6th Framework Programme IST Multimodal Interfaces, http://emotion-research.net. The National Centre of Competence in Research (NCCR) in Affective Sciences financed by the Swiss National Science Foundation (n° 51NF40-104897). A grant from the Swiss National Science Foundation (105311-108187/1 to David Sander and Patrik Vuilleumier). The “Programme d’actions intégrées Franco-Suisse Germaine de Staël” in collaboration with the Swiss Academy for Technical Sciences (to Lionel Reveret and David Sander).The authors would like to thank Prof. Susanne Kaiser, Dr. Marc Méhu, Katia Schenkel, Birgit Michel, and Stéphane With (University of Geneva) for their expertise and guidance about the FACS, and Dr Mina Vasalou (University of Bath) for comments on drafts of this paper.

References

Bickel, B., Botsch, M., Angst, R., Matusik, W., Otaduy, M., Pfister, H., et al. (2007). Multi-scale capture of facial geometry and motion. *ACM Transactions in Graphics, 26*(3), 33.

Blanz, V., & Vetter, T. (1999). *Morphable model for the synthesis of 3D faces [Computer software]*. Los Angeles: SIGGRAPH.

Corneille, O., Hugenberg, K., & Timothy, P. (2007). Applying the attractor field model to social cognition: Perceptual discrimination is facilitated, but memory is impaired for faces displaying evaluatively congruent expressions. *Journal of Personality and Social Psychology, 93*(3), 335–352.

Cosker, D., Borkett, R., Mashall, D., & Rosin, P. L. (2008). Towards automatic performance driven animation between multiple types of facial model. *IET Computer Vision, 2*(3), 129–141.

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

Delplanque, S., N'Diaye, K., Scherer, K. R., & Grandjean, D. (2007). Spatial frequencies or emotional effects? A systematic measure of spatial frequencies for IAPS pictures by a discrete wavelet analysis. *Journal of Neuroscience Methods, 165*, 144–150.

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

Ekman, P., Friesen, W. V., & Hager, J. (2002). *The facial action coding system.* London, UK.

Ekman, P., Irwin, W., & Rosenberg, E. L. (1994). *The emotional facial action coding system (EMFACS).* London, UK.

Ellison, J. W., & Massaro, D. W. (1997). Featural evaluation, integration, and judgment of facial affect. *Journal of Experimental Psychology: Human Perception and Performance, 23*(1), 213–226.

FaceGen Modeller. [Software] (2009). Singular Inversions Inc. Retrieved from http://www.facegen.com/.

Fiser, J., Bex, P. J., & Makous, W. (2003). Contrast conservation in human vision. *Vision Research, 43*, 2637–2648.

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

Gaag, C., van der Minderaa, R. B., & Keysers, C. (2007). The bold signal in the amygdala does not differentiate between dynamic facial expressions. *Social Cognitive Affective Neuroscience, 2*, 93–103.

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: A validation studies. *Cognition and Emotion, 22*(6), 1094–1118.

Grammer, K., Tessarek, A., & Hofer, G. (in preparation). From emoticons to avatars: The simulation of facial expression. In A. Kappas (Ed.), *Emotional communication on the internet*. Retrieved from http://evolution.anthro.univie.ac.at/institutes/urbanethology/projects/simulation/emosym/index.html.

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior, 21*(4), 241–257.

Hezle, V., Biehn, C., Schlömer, T., & Linner, F. (2004). Adaptable setup for performance driven facial animation. In *Proceedings of SIGGRAPH'04—sketches.* Los Angeles: Springer.

Hirsh, A. T., Alqudah, A. F., Stutts, L. A., & Robinson, M. E. (2009). Virtual human technology: Capturing sex, race, and age influences in individual pain decision policies. *Pain, 140*, 231–238.

Jooman, J., & Gotlib, I. H. (2006). Is this happiness I see? Biases in the identification of emotional facial expressions in depression and social phobia. *Journal of Abnormal Psychology, 115*(4), 705–714.

Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. In *Proceedings of the fourth IEEE international conference on automatic face and gesture recognition.* Retrieved from http://vasc.ri.cmu.edu/idb/html/face/facial_expression/index.html.

Lundqvist, D., Esteves, F., & Öhman, A. (1998). The Karolinska directed emotional faces—KDEF [CD ROM]. Department of Clinical Neuroscience, Psychology section, Karolinska Institute. ISBN 91-630-7164-9.

Ma, W., Jones, A., Chiang, J., Hawkins, T., Frederiksen, S., Peers, P., et al. (2008). Facial performance synthesis using deformation-driven polynomial displacement maps. *ACM Transactions in Graphics, 27*(5), 1–10.

Malatesta, L., Raouzaiou, A., Karpouzis, K., & Kollias, S. (2006). Mpeg-4 facial expression synthesis based on appraisal theory. In The *3rd IFIP conference in artificial intelligence applications and innovations, AIAI 2006*. Athens, Greece.

Moradi, F., Koch, C., Shimojo, S., Sarma, G., & Gutierrez, J. (2005). Adaptation to face identity and emotional expression depends on attention. In *Proceedings of vision sciences society 5th.* Sarasota, FL: Journal of Vision.

Moser, E., Derntl, B., Robinson, S., Fink, B., Gur, R. C., & Grammer, K. (2007). Amygdala activation at 3t in response to human and avatar facial expressions of emotions. *Journal of Neuroscience Methods, 161*(1), 126–133.

N'Diaye, K., Sander, D., & Vuilleumier, P. (2009). Self-relevance processing in the human amygdala: Gaze direction, facial expression, and emotion intensity. *Emotion, 9*(6), 798–806.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. In *Proceedings of the national academy of sciences of the United States of America, 105*(32), 11087–11092.

Pantic, M., Valstar, M. F., Rademaker, R., & Maat, L. (2005). Web-based database for facial expression analysis. In *Proceedings of the IEEE international conference on multimedia and expo (ICME'05)*. Retrieved from http://www.docstoc.com/docs/2933918/EULA-End-User-License-Agreement-MMI-Face-Database-www-mmifacedb.

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

Parr, L. A., Waller, B. M., & Heintz, M. (2008). Facial expression categorization by chimpanzees using standardized stimuli. *Emotion, 8*(2), 216–231.

Pasquariello, S., & Pelachaud, C. (2001). Greta: A simple facial animation engine. In R. Rajkumar, M. Köppen, S. Ovaska, T. Furuhashi, & F. Hoffmann (Eds.), *6th online world conference on soft computing in industrial applications, session on soft computing for intelligent 3D agents*. Germany: Springer.

Pelphrey, K., Viola, R., & McCarthy, G. (2004). When strangers pass: Processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological Science, 15*(9), 598–603.

Poser 7 [Software] (2007). e-frontier. Retrieved from http://www.e-frontier.com/go/products/poser/.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex, 14*(6), 619–633.

Roesch, E. B., Sander, D., Mumenthaler, C., Kerzel, D., & Scherer, K. R. (2010a). Psychophysics of emotion: The QUEST for emotional attention. *Journal of Vision, 10*(3), 4, 1–9.

Roesch, E. B., Sander, D., & Scherer, K. R. (2009). Emotion and motion in facial expressions modulate the attentional blink. *Perception, 38*, 466.

Roesch, E. B., Sander, D., & Scherer, K. R. (2010b). *The 4th dimension(s) of emotion perception: Emotion and motion in facial expressions modulate the attentional blink.* Manuscript in preparation.

Ruys, K. I., & Stapel, D. A. (2008). Emotion elicitor or emotion messenger: Subliminal priming reveals two faces of facial expressions. *Psychological Science, 19*(6), 593–600.

Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., & Scherer, K. R. (2007). Interaction effect of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *European Journal of Cognitive Psychology, 19*(3), 470–480.

Sayette, M. A., Cohn, J. F., Wertz, J. M., Perrott, M. A., & Parrott, D. J. (2004). A psychometric evaluation of the facial action coding system for assessing spontaneous expression. *Journal of Nonverbal Behavior, 25*(3), 167–186.

Scherer, K. R., & Ellgring, H. (2007). Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion, 7*(1), 113–130.

Schulte-Rüther, M., Markowitsch, H. J., Fink, G. R., & Piefke, M. (2007). Mirror neuron and theory of mind mechanisms involved in face-to-face interactions: A functional magnetic resonance imaging approach to empathy. *Journal of Cognitive Neuroscience, 19*, 1354–1372.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6*(12), 1317–1322.

Szczepanowski, R., & Pessoa, L. (2007). Fear perception: Can objective and subjective awareness measures be dissociated? *Journal of Vision, 7*(4), 10–17.

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social Cognitive Affective Neuroscience, 3*, 119–127.

Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology, 78*(1), 105–119.

Zhang, L., Snavely, N., Curless, B., & Seitz, S. (2004). Spacetime faces: High resolution capture for modeling and animation. *ACM Transactions in Graphics, 23*(3), 548–558.