

Making faces: Creating three-dimensional parameterized models of facial expression

JESSE SPENCER-SMITH, HEATHER WILD, ÅSE H. INNES-KER, JAMES TOWNSEND,
CHRISTY DUFFY, CHAD EDWARDS, KRISTINA ERVIN, NICOLE MERRITT, and JAE WON PAIK
Indiana University, Bloomington, Indiana

Previous studies of the perceptual processing and dynamics of emotional expression in faces have been limited by the lack of realistic yet controlled stimuli. The present work offers researchers a method for creating such stimuli for exploring these phenomena. We describe the creation of the stimuli and a series of experiments testing the validity of these stimuli with respect to emotional expressions in humans. Participants evaluated synthesized facial images and standardized photographs of six basic emotional expressions for intensity and accuracy of perceived emotion. Comparisons of these measures were qualitatively similar for synthesized and photographed faces. A manipulation of the magnitude of the synthesized expressions yielded a significant effect on the perceived intensity of expression. In a subsequent multidimensional scaling study, no systematic differences were uncovered in the derived configurations of the synthesized expressions and the photographs. These results are discussed in the context of possible future research applications.

When a Spaniard smiles at an Aboriginal Australian, or when a Yanomamo Indian scowls at a Brazilian, the perceiver knows the emotional import of the facial expression almost immediately. When a person you are interacting with shows a fleeting smile, you feel encouraged to continue the interaction without change; if the facial expression drops toward a frown, you make rapid adjustments in what you are doing in order to address and perhaps alleviate the sudden displeasure. A face breaking into a smile signals joyous happiness, whereas a smile slowly creeping upon a face expresses dawning understanding.

In the present study, we seek to offer researchers a method for creating controlled and ecologically valid stimuli for exploring these phenomena. The paper is organized as follows. First, we describe research on modeling emotional expressions in faces, including the approaches that have arisen in computer animation for modeling such expressions. Second, we describe in depth how we created the stimuli used in the present study. Third, we describe a series of experiments testing the validity of these stimuli with respect to emotional expressions in humans. Finally, we evaluate and discuss the results in terms of possible contributions to future research.

Much of our understanding of human expressions has come from research using high-intensity, unambiguous emotional expressions presented via still photographs. Research over the past four decades suggests that six basic

emotional expressions—anger, disgust, fear, happiness, sadness, and surprise—are recognized reliably (e.g., Ekman, Friesen, & Ellsworth, 1982; Russell, 1994), universally (e.g., Ekman, 1971; but see Russell, 1994), and rapidly (Kirouac & Dore, 1984; McAndrew, 1986). However, the perceptual and decisional processes underlying these findings are poorly understood. Also, work with low-intensity expressions has not been as prevalent (Hess, Blairy, & Kleck, 1997), although our everyday experience with subtle expressions suggests that we are similarly adept at reading subtle emotions. Likewise, the dynamics of expressions, although widely recognized as important to our understanding of emotional expression, has received relatively limited attention (Niedenthal, Halberstadt, Margolin, & Innes-Ker, 2000; Spencer-Smith, Innes-Ker, & Townsend, 2000; Wehrle, Kaiser, Schmidt, & Scherer, 2000). Research in all of areas has historically suffered from a lack of stimuli.

The purpose of this study is to create a three-dimensional, flexible, parameterized, ecologically motivated model of facial expressions. We seek to incorporate the ability to blend movements and expressions, to produce meaningful low-intensity emotions, to allow for the creation of dynamic stimuli, and to validate the model against standard static images of expressions.

Much research on emotional expressions has used image morphing (e.g., Calder, Young, Perrett, & Etcoff, 1996), creating transitional images by using images of facial expressions with high emotional intensity. The use of such techniques presents several difficulties. Whereas the original images from which the morphs are derived are of humans expressing emotions, the in-between, or morphed, images do not necessarily represent physiologically possible expressions. In addition, Pittenger (1991), in a critique of image-morphing approaches, demonstrated that

Funding for this project was provided by the Indiana University President's Summer Undergraduate Research Initiative and by NIH-NIMH Grant R01MH57717. Correspondence concerning this article should be addressed to J. Spencer-Smith, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 405 North Mathews Avenue, Urbana, IL 61801 (e-mail: jbspence@beckman.uiuc.edu).

image morphing of two three-dimensional objects does not necessarily produce a veridical average of the objects. In recent years, alternative three-dimensional approaches have been investigated, providing varying degrees of physiological and ecological validity.

One approach involves three-dimensional capture of human expressions. One such technique uses highly reflective markers on key points on an actor's face, such as the tip of the nose, the tips and midpoints of the lips and brows, the cheeks, and so forth. When lit with an ultraviolet light, the markers reflect back light that can be detected by cameras positioned directly in front and to the side. A computer then calculates the position of each marker in three-dimensional space. Each position is then sent to a computer-generated model of the face, and corresponding points of the model that represent each marker are moved accordingly (Guenther, Grimm, Malvar, & Wood, 1998). Although such a technique is simple to implement and quick in calculation, it lacks fine definition. The method also does not allow for the creation of novel expressions not posed for by the actor.

Another approach taken is the modeling of each facial muscle onto a modeled human skull. For each expression, appropriate muscle shape can be altered in the computer environment to simulate the real muscle being flexed. Superimposed skin is then moved with the muscles, creating an expression. Although this method is physiologically sound, it is inflexible in that it is difficult to model multiple identities. New faces have to be constructed from the skull, to the muscles, and then up to the skin. Such tasks require an expert in facial anatomy and reconstruction techniques.

One of the most recently developed methods is the photogrammetric technique in which several simultaneous profile shots are taken of an actor from different profiles. A three-dimensional image is calculated by using the flat images. Skin is then texture-mapped onto the figure, allowing small wrinkles and crevices to be easily simulated. Three-dimensional morphing animates expressions (Pighin, Szeliski, & Salesin, 1999). Although this morphing technique is used most often in the facial animation industry, it still requires trained actors. And like the earlier method using actors, there is no systematic way to generate novel images or to exercise fine control over the intensity of expressions.

The basis for our model is a three-dimensional mesh-based character animation package that allows for a great deal of control in conducting research into the dynamic nature of expression recognition without sacrificing ecological validity. The stimuli would also be invaluable for probing the perception of static emotional expressions, since they easily allow for the necessary fine-grained control over stimuli.

Because it is of the utmost importance that the stimuli reflect as closely as possible the wide range of expressiveness of humans, we opted for an approach for creating the stimuli in which expressions are built, bottom up, from the identified muscular elements of facial expressiveness.

The approach is similar to that of Wehrle et al. (2000). However, we created full, realistic three-dimensional images, whereas Wehrle et al. used two-dimensional line drawings.

As a reference for the stimulus creation, we relied on Ekman and Friesen's (1976a) Facial Action Coding System (FACS). The express purpose of FACS was to create a coding system for precisely cataloging facial movements in naturally occurring faces without having to resort to imprecise or emotion-related terminology. The basic element of the system is the action unit (AU). An AU is defined as the surface (skin) structure resulting from the action of a facial muscle group that can be controlled independently from all other facial muscle groups. Expressions can be coded by detecting the presence of combinations of AUs in the face. The FACS training manual provides photographs, movie clips, precise written descriptions, and a description of the musculature implicated in each AU.

In the present study, we implemented 16 AUs on simulated female and male faces, using a three-dimensional mesh-based modeling program, Poser 4 (Curious Labs, Santa Cruz, CA). Poser provides realistic default female and male characters defined by polygons (3,470 vertices for the female character head, 3,427 for the male). Using the software, one can render images of full-body characters from arbitrary views, wrap new skin textures around faces, and arbitrarily define lighting. New characters can be created by distorting default characters by means of morph targets; morph targets describe three-dimensional deformations to the default geometry. One morph target, for example, alters the length of the face of the character by moving down the vertices of several hundred polygons defining the face. New morph targets can be created by manually moving vertices, using modeling tools provided in the program. These distortions can then be used in concert with other morph targets to create new effects, such as a long, wide face. The distortions are applied to the face by means of a parameter setting, one for each morph target. For example, at a magnitude setting of 1.0, a user-defined morph target will recreate the original distortion defined by the user; at 0.5, the deformation is less severe and is closer to the default face. At a setting of 2.0, a morph target exaggerates the deformation.

Using Ekman and Friesen's (1976a) FACS as a guide, our group implemented 16 of the 44 AUs in the FACS training manual as morph targets. To illustrate, we will outline the process of implementing AU 24, the lip presser, for the Poser male face, as can be seen in Figure 1. AU 24 presses the lips together, without moving the chin boss (the rounded front part of the chin), tightens and narrows the lips, and creates a bulging of the skin above the upper lip and/or below the lower lip. Ekman and Friesen (1976a) provide a detailed account of the muscles implicated in the movement, a description of the appearance and how it might vary between individuals, multiple photographs of actors posing AU 24, as well as movie clips of the action being performed. One member of our team took the



Figure 1. Poser male character with AU 24 morph target set to 0 and to 1.0.

initial responsibility of manually altering a default Poser male face, using the software modeling tools to recreate AU 24 on the male face. This involved selecting and manually moving vertices in three dimensions, in a process that could be characterized as virtual sculpting. Using multiple views, the team member worked over several days until the result matched the pictorial and the written descriptions. The full team then critiqued the AU, using multiple views and lighting conditions. If necessary, additional changes were made until the morph target closely matched the FACS description. At this point, AU 24 then could be added to other male characters or could be combined with other AUs to create complete expressions. To create a low-magnitude AU 24, representing a minor lip pressing, the AU 24 morph target parameter could be set to less than 1. To create extreme lip pressing, the parameter could be set to greater than 1. Finally, AU 24 was added to a male figure with all of the AUs to check that it would interact with other AUs in the manner prescribed by the FACS training manual.

With each AU implemented as an independently controllable morph target, it was possible to create novel expressions, to create low-magnitude expressions, and to generate dynamic displays of expressions. The models offer fine temporal control over the rise and fall of the activations for each individual AU. Multiple identities can be created by using other morph targets to alter facial structure, while allowing identical expressions to be created on each character by setting equivalent AU morph target parameter settings. Lighting and viewpoint can also be set arbitrarily.

The stated goal for the project was to create a flexible, parameterized, ecologically motivated model of facial movement with the ability to blend movements and expressions, to produce meaningful low-intensity emotions, and to allow for the creation of dynamic stimuli. To validate our morph target AUs, in three experiments, we generated versions of the six basic expressions—anger, disgust, fear, happiness, sadness, and surprise—and compared the resultant images with Ekman and Friesen's (1976b) pictures of facial affect. In the first study, the participants rated the emotional intensity and accuracy of expressions of synthesized female characters and of the photographs. To test our ability to control perceived emotional intensity, we included expressions created by using

both high-magnitude and low-magnitude AU morph targets. In the second study, we studied synthesized male characters, using the same procedure. In the third study, we paired images of synthesized male and female characters with the photographs in a dissimilarity ratings task; we then submitted these data to multidimensional scaling (MDS) routines and compared the resulting configurations.

If the models of expressive faces created during this project were sufficiently realistic, we would expect to find similar ratings of intensity and accuracy for the full-magnitude expressions and photographs. In addition, we would hope to find similar patterns and confusions in the intensity ratings between the Poser stimuli, the photographs, and the published reports. If we were successful in manipulating intensity by using magnitude settings, we would expect to find significantly lower ratings of intensity for low-magnitude expressions. Finally, we would expect that the Poser images and the photographic stimuli of specific expressions would lie close to one another in an MDS space derived from the dissimilarity measures.

EXPERIMENT 1

Method

Participants. Twenty participants (7 males and 13 females) completed Experiment 1. All the participants were either undergraduate or graduate students at Indiana University, and all were paid for their participation.

Stimuli. The stimuli consisted of 119 grayscale images of faces. Seventy-seven images were photographs of faces taken from Ekman and Friesen's (1976b) *Pictures of Facial Affect*. There were photographs of 11 individuals (five males and six females) with seven expressions each. Forty-two images were generated using Poser (see the description of stimulus creation above). There were six synthesized Poser characters, female in appearance. Each character was rendered with seven different emotional expressions—anger, disgust, fear, happiness, sadness, surprise, and neutral—using the AUs listed in Table 1 (see Figure 2). Two levels of AU morph target magnitude were used in rendering the computer-generated images: half of the computer-generated faces were designed with magnitude settings of approximately 0.5, and the other half were designed with magnitude settings close to 1.0 (see Figure 3 for a low-magnitude example). The Poser figures were rendered in grayscale with viewpoint, camera focal length, and lighting to match the Ekman and Friesen (1976b) photographs as closely as possible. The images and photographs measured 320 x 480 pixels.

Procedure. There were two parts to this experiment: an intensity-rating task and an accuracy-rating task. The participants were read a short set of instructions explaining that the goal of the experiment was to test how people perceive emotional expressions in faces. The participants were told that some of the images would be computer

Table 1
Component Action Units for Expressions
(Parke & Waters, 1996)

Expression	Action Unit
Surprise	1+2
Fear	1+2+4+15/16+20
Disgust	4+9+10+17
Anger	2+4+5+10+24/25
Happiness	6+11+12
Sadness	1+2+4+15

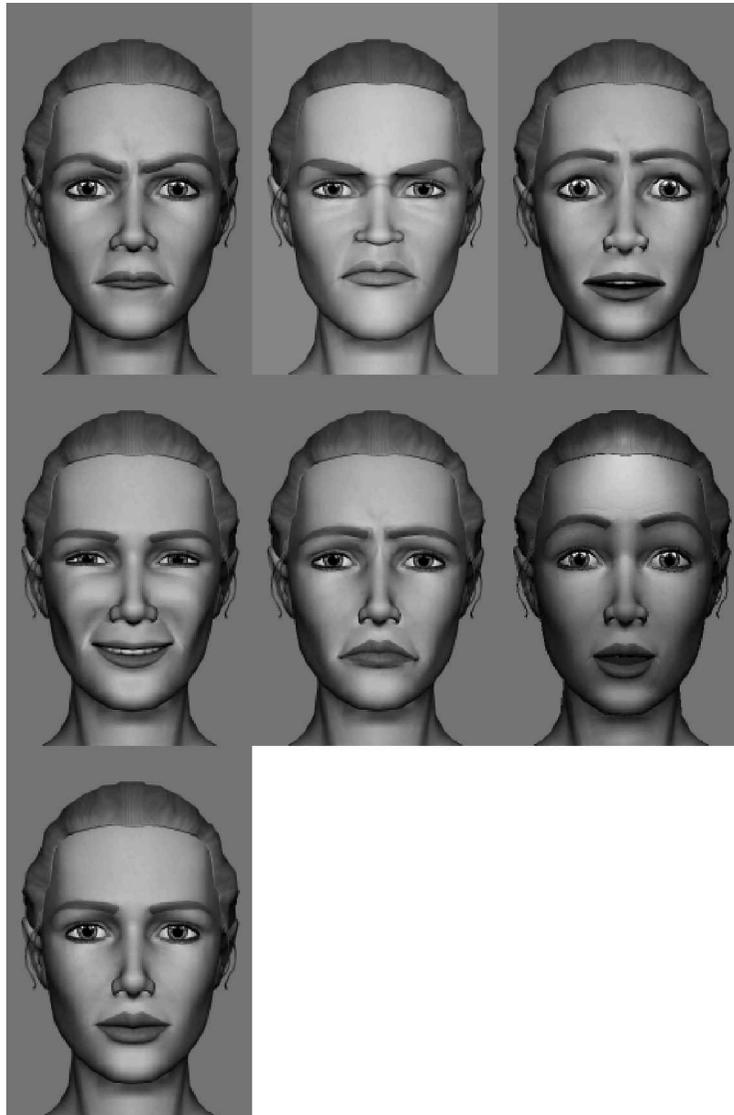


Figure 2. Poser female characters with expressions of anger, disgust, fear, happiness, sadness, surprise, and neutrality, used in Experiment 1.

generated, whereas others would be photographs of faces. They were reminded that neither format is necessarily better at depicting any emotion and were told to follow their first and immediate impression. They were asked to answer quickly and spontaneously and to use the entire scale.

For the intensity-rating task, each participant saw all 119 images (77 photographs and 42 computer-generated images). Images were presented in random order. The participants rated each image for each of the seven emotions (anger, disgust, fear, happiness, sadness, surprise, neutral). Question order was randomized across participants.

The participants viewed faces one at a time, with unlimited viewing time, and rated how intensely each emotion was expressed by the face, using a 7-point Likert scale in which 1 equaled *no intensity* and 7 equaled *high intensity*. The scale was presented below each image. After the participant responded, another question appeared inquiring about another emotion. After the participants rated the face on each of seven emotions, another image appeared, and the sequence was repeated. The participants took a short break at the

halfway point. When the participants completed the intensity-rating task, they began the accuracy-rating task.

In the accuracy-rating part of the experiment, the participants were asked to rate how accurately a particular expression conveyed its intended expression. In addition to the images detailed earlier, the participants viewed six additional Poser figures designed to express emotions with both high and low intensities. The participants were asked to use a 7-point scale, ranging from *very poorly* to *very accurately*.

Results and Discussion

The participants' responses for the first part of Experiment 1 are summarized in Table 2. Responses to each of the 11 actors expressing anger, disgust, fear, happiness, sadness, and surprise are averaged across the 20 participants in the top half of the table. Responses to each of the three Poser characters with the same expressions com-



Figure 3. Sample Poser female figure with expression of happiness composed of low-magnitude action units.

posed of high-magnitude AUs are summarized in the lower half. It is evident that, like the photographs, the Poser images convey multiple expressions. In expressions intended to convey happiness, for example, the participants rated surprise as the second strongest emotion in both the photographs and the Poser images. Overall, combining AU morph targets successfully communicated the intended emotional content. With two exceptions, the emotion rated as the most intense was the intended emotion. In the first exception, anger was rated as the most intense emotion in an expression intended to convey disgust. The distinguishing AU for disgust is AU 9, which involves a wrinkling at the base of the nose. Low polygon

counts in the nasal region made clear depiction of this wrinkling difficult. We note that, in the photographs, the second highest rated emotion for the disgust expression was anger. The second exception occurred with the expression intended to convey fear. In this expression, surprise was rated as the most intense emotion. Surprise is a valence-neutral expression, meaning that it is not necessarily a positive or a negative expression. It consists of AUs 1 and 2, which are present in fear and in sadness. The high rating of surprise for this expression can be attributed to an overly high setting of these AU morph targets.

Average judged intensity varied significantly with image type [photograph, Poser with high-magnitude AU morph targets, Poser with low-magnitude morph targets; $F(2,18) = 49.02, p < .0001$]. Planned contrasts revealed that high-magnitude Poser images were rated as less intense than the photographs [$M = 5.732, SE = 0.086; F(1,19) = 5.449, p < .031$]. High-magnitude Poser images ($M = 5.469, SE = 0.143$) were judged as more intense than low-magnitude Poser images [$M = 4.543, SE = 0.150; F(1,19) = 85.0, p < .0001$].

Although the intensities of the Poser high-magnitude images were judged as less intense overall than the photographs, two emotions, disgust and happiness, were primarily responsible for the difference. On the other hand, emotional intensity was successfully manipulated by reducing the magnitude of the component AU morph targets. Presumably, the magnitude of the AU morph targets making up disgust and happiness could be raised to increase the perceived emotional intensity. This was attempted in the study of the male Poser characters in Experiment 2.

Average judged accuracy varied significantly with image type [photograph, Poser with high-magnitude AU morph targets, Poser with low-magnitude morph targets; $F(2,17) = 17.83, p < .0001$]. Planned contrasts revealed that high-magnitude Poser images ($M = 5.664, SE =$

Table 2
Intensity Ratings for Ekman and Friesen (1976b) Photographs and High-Magnitude Poser Female Character Expressions

Expression	Ratings						
	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Neutral
Photographs							
Anger	5.64	4.07	2.43	1.26	2.28	2.44	1.55
Disgust	4.06	6.12	1.81	1.27	1.97	1.73	1.57
Fear	2.61	2.80	5.38	1.39	2.45	5.35	1.46
Happiness	1.19	1.10	1.12	6.45	1.17	2.37	1.80
Sadness	2.51	2.81	2.73	1.18	5.49	1.71	2.00
Surprise	1.77	2.01	3.30	2.16	1.74	6.25	1.65
Neutral	2.51	2.40	1.83	1.79	2.70	1.43	4.79
Poser Images							
Anger	5.22	4.10	3.12	1.20	3.57	2.13	1.58
Disgust	5.97	4.70	1.77	1.20	2.75	1.50	1.58
Fear	2.45	2.72	4.90	1.35	3.78	5.38	1.55
Happiness	1.27	1.22	1.33	6.00	1.27	2.38	2.03
Sadness	3.05	2.83	2.67	1.35	6.00	1.73	2.02
Surprise	1.40	1.45	2.73	3.15	1.75	6.27	1.75
Neutral	1.88	1.83	2.10	2.28	2.83	1.50	5.20

Note—Boldface indicates intensity rating of intended emotion.



Figure 4. Sample Poser male characters with expressions of anger, disgust, fear, happiness, sadness, surprise, and neutrality, used in Experiment 2.

0.111) were judged as less accurate than the photographs [$M = 5.904$, $SE = 0.122$; $F(1,18) = 8.124$, $p < .011$]. The difference, however, was less than 0.3 on a scale from 1 to 7.

EXPERIMENT 2

Method

Participants. Twenty participants (7 males and 13 females) completed Experiment 2. All the participants were either undergraduate or graduate students at Indiana University, and all were paid for their participation. None of the participants had participated in Experiment 1.

Stimuli. The same photographs were used as those in Experiment 1. The synthetic stimuli were Poser male faces. Since the default geometry for the Poser female and male faces contains a different number of polygons, new AU morph targets had to be created for the Poser male faces. Examples of these stimuli appear in Figure 4.

Procedure. The procedure was identical to that in Experiment 1.

Results and Discussion

The participants' responses for the intensity-rating portion of Experiment 2 are summarized in Table 3. With a single exception, the emotion rated as most intense was the intended emotion. As in Experiment 1, anger was rated as the most intense emotion in an expression intended to convey disgust. Without wrinkles to indicate movement of skin along the base of the nose, AU 9 remains difficult to capture with the current geometry. By manipulating AU morph targets 1 and 2 in the fear expression, we were able to correct the confusion between fear and surprise found in Experiment 1.

Average judged intensity varied significantly with image type [photograph, Poser with high-magnitude AU morph targets, Poser with low-magnitude morph targets;

Table 3
Intensity Ratings for Ekman and Friesen (1976b) Photographs and High-Magnitude Poser Male Character Expressions

Expression	Ratings						
	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Neutral
Photographs							
Anger	5.45	3.73	2.25	1.19	2.21	2.32	1.50
Disgust	3.90	5.78	1.81	1.23	2.12	1.91	1.39
Fear	2.07	2.63	5.54	1.22	2.36	5.40	1.30
Happiness	1.15	1.15	1.15	6.40	1.19	2.14	1.69
Sadness	2.29	2.40	2.42	1.15	5.26	1.46	2.20
Surprise	1.69	1.95	3.59	2.11	1.60	6.10	1.47
Neutral	2.06	2.02	1.51	1.75	2.66	1.36	5.59
Poser Images							
Anger	6.37	4.43	1.93	1.13	2.43	1.70	1.32
Disgust	5.22	4.80	1.92	1.15	2.67	1.67	1.42
Fear	1.67	2.13	5.23	1.40	2.88	5.07	1.68
Happiness	1.08	1.05	1.18	6.40	1.23	2.12	1.43
Sadness	1.62	1.88	2.35	1.07	5.97	1.53	2.30
Surprise	1.17	1.47	2.67	3.50	1.58	5.53	1.87
Neutral	1.47	1.47	1.37	3.27	2.18	1.32	5.63

Note—Boldface indicates intensity rating of intended emotion.

$F(2,18) = 57.15, p < .0001$]. Planned contrasts revealed that high-magnitude Poser images did not have a significantly different intensity from the photographs [$M = 5.733, SE = 0.106; F(1,19) = 0.094, p > .828$]. High-magnitude Poser images ($M = 5.705, SE = 0.139$) were again judged as more intense than low-magnitude Poser images [$M = 4.88, SE = 0.150, F(1,19) = 91.33, p < .0001$].

Average judged accuracy varied significantly with image type [photograph, Poser with high-magnitude AU morph targets, Poser with low-magnitude morph targets; $F(2,19) = 48.04, p < .0001$]. Planned contrasts revealed that high-magnitude Poser images ($M = 5.69, SE = 0.109$) were not judged as less accurate than the photographs, as was found in Experiment 1 [$M = 5.554, SE = 0.131$;

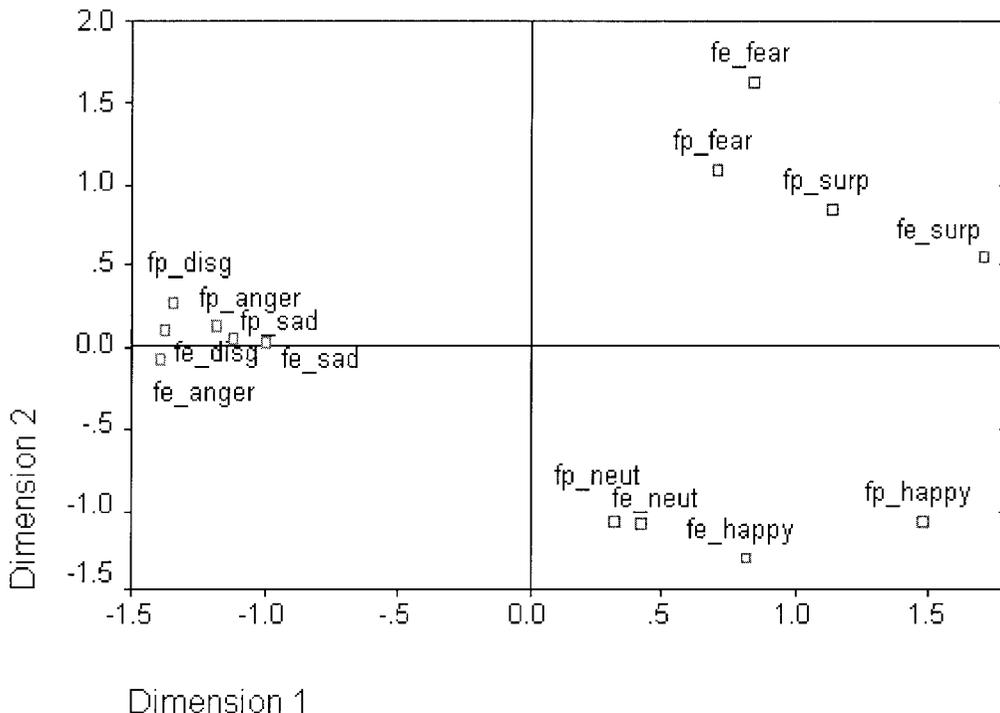


Figure 5. Derived two-dimensional multidimensional scaling configuration for Poser female and Ekman and Friesen (1976b) faces. “fe” corresponds to Ekman and Friesen faces, and “fp” corresponds to Poser faces.

$F(1,20) = 2.90, p = .104$]; the trend was for the Poser images to be rated as more accurate than the photographs.

EXPERIMENT 3

Method

Participants. Nine participants (5 males and 4 females) completed Experiment 3. All the participants were either undergraduate or graduate students at Indiana University and were paid for their participation. None had participated in Experiment 1 or Experiment 2.

Stimuli. The stimuli consisted of 28 grayscale images of emotionally expressive faces. Half of these images were generated using Poser, whereas the other half were photographs (Ekman & Friesen, 1976b). There was one synthesized male and one synthesized female character, each rendered with seven different emotional expressions (anger, fear, disgust, happiness, sadness, surprise, and neutral). The photographs were of a male and a female actor posing the same expressions. The images measured 320×480 pixels.

Procedure. The participants were read a short set of instructions explaining that some of the images would be computer-generated faces, whereas others would be photographs of faces. They were asked to judge the dissimilarity of two expressions, regardless of the nature of the images. The participants were also presented with the same set of instructions on the computer screen.

The participants viewed randomly paired combinations of faces. The gender of the faces was blocked, with the participants receiving either male or female faces first, on the basis of random assignment. The participants rated all possible pairs of faces. Two images appeared on the screen side by side. Below the images, instructions appeared indicating that the participants should lengthen or shorten a bar to signify how dissimilar the expressions were. The participants lengthened and shortened the bar by holding down the \ or the backspace key, respectively. If the expressions in the two images were very different, they were instructed to make the bar very long. If the expressions in the two images were almost identical, they were instructed to make the bar very short. At its longest, the bar

contained 216 backlash characters. The participants had 14 practice trials before beginning each half of the experiment.

Results and Discussion

The dissimilarity ratings for the Poser and the Ekman and Friesen (1976b) faces were submitted to a three-way¹ scaling MDS analysis using ALSCAL three-way analysis (INDSCAL algorithm). The MDS model yields a spatial representation of the proximity data (in this case, dissimilarity ratings), wherein each object is represented by a point in multidimensional space and the distance between points is inversely proportional to the similarity of the objects. Different directions in the resultant configuration can often be interpreted as corresponding to dimensions that are the basis of observers' proximity ratings.

A separate two-dimensional MDS analysis was performed for the ratings for the male and the female faces. The resultant configuration for the female faces appears in Figure 5. The stress value² for this solution was .15. The resultant configuration for the male faces appears in Figure 6; the stress value for this solution was .13.

The resultant MDS configurations yield a visual representation of how similar the Ekman and Friesen (1976b) faces are to the Poser faces that is based on the dissimilarity ratings given by our observers. Many of the data points for corresponding emotions are fairly close together. The Poser disgust is noticeably distant from the Ekman and Friesen disgust; we anticipated this result on the basis of the high rating of anger on the Poser disgust image in the previous experiments. Although the configurations are noisy, especially in the case of the Poser female dissimilarity ratings, inspection revealed no system-

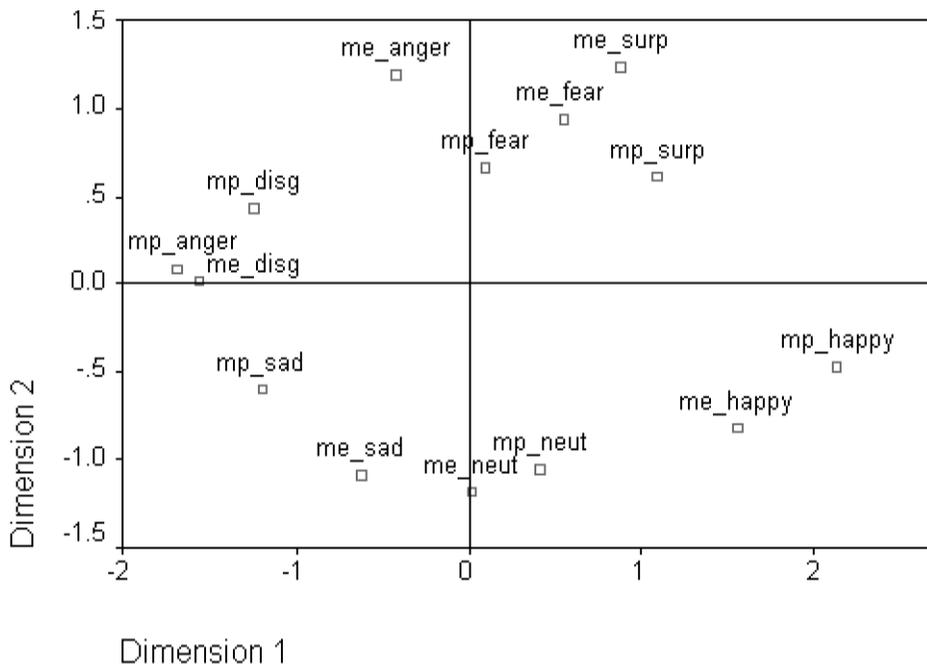


Figure 6. Derived two-dimensional multidimensional scaling configuration for Poser male and Ekman and Friesen (1976b) faces; "me" corresponds to Ekman and Friesen faces and "mp" corresponds to Poser faces.

atic deviations between the Poser images and the Ekman and Friesen images. The configurations did not change qualitatively when the Poser and photograph data were fit separately, suggesting that the Poser images shared the same dimensions as the photographs.

GENERAL DISCUSSION

In this study, we implemented Ekman and Friesen's (1976a) AUs in a three-dimensional mesh-based character animation program. The AUs were instantiated as morph targets, three-dimensional deformations of standard synthesized male and female faces. Although the mesh model itself is not based on musculature, by using AUs as a guide, we ensured that the motions produced by individual morph targets are feasible, realistic facial motions. The AU morph targets are flexible: They can be applied alone or in groups to create expressions and can be used to generate novel expressions. The AU morph targets are parameterized and can be applied with a specified magnitude. When used in generating an expression, we found that we could manipulate perceived intensity of the expression by adjusting the magnitudes of the component AUs. By basing our AU morph targets on well-researched observations of facial movements, the morph targets are ecologically based. Most important, the AU morph targets allow for the creation of dynamic expression stimuli in a controlled manner.

The successful manipulation of intensity will allow investigations into low-intensity expressions that have received scant attention. Spencer-Smith et al. (2000), using the AU morph targets developed in this project, have demonstrated the critical role of dynamics in perceiving low-intensity emotional expressions. Very low emotional intensity expressions were accurately identified when presented in a dynamic fashion, whereas accuracy with static displays was near chance.

REFERENCES

- CALDER, A. J., YOUNG, A. W., PERRETT, D. I., & ETCOFF, N. L. (1996). Categorical perception of morphed facial expressions. *Visual Cognition*, *3*, 81-117.
- EKMAN, P. (1971). Universality and cultural differences in facial ex-

- pressions. In J. K. Cole (Ed.), *Nebraska Symposium on Motivation* (Vol. 19, pp. 207-283). Lincoln: University of Nebraska Press.
- EKMAN, P., & FRIESEN, W. V. (1976a). Measuring facial movement. *Environmental Psychology & Nonverbal Behavior*, *1*, 56-75.
- EKMAN, P., & FRIESEN, W. V. (1976b). *Pictures of facial affect*. Unpublished manuscript.
- EKMAN, P., FRIESEN, W. V., & ELLSWORTH, P. (1982). What are the relative contributions of facial behavior and contextual information to the judgment of emotion? In P. Ekman (Ed.), *Emotion in the human face* (2nd ed. pp. 111-127). New York: Cambridge University Press.
- GUENTER, B., GRIMM, C., MALVAR, H., & WOOD, D. (1998). Making faces. In *SIGGRAPH 98 conference proceedings*. New York: ACM Press.
- HESS, U., BLAIRY, S., & KLECK, R. E. (1997). The intensity of emotional facial expression and decoding accuracy. *Journal of Nonverbal Behavior*, *21*, 241-257.
- KIROUAC, G., & DORE, F. Y. (1984). Judgment of facial expressions of emotion as a function of exposure time. *Perceptual & Motor Skills*, *59*, 147-150.
- MCANDREW, F. T. (1986). A cross-cultural study of recognition thresholds for facial expressions of emotion. *Journal of Cross-Cultural Psychology*, *17*, 211-224.
- NIEDENTHAL, P. M., HALBERSTADT, J. B., MARGOLIN, J., & INNES-KER, A. H. (2000). Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology*, *30*, 211-222.
- PARKE, F. I., & WATERS, K. (1996). *Computer facial animation*. Wellesley, MA: Peters.
- PIGHIN, F., SZELISKI, R., & SALESIN, D. (1999). Resynthesizing facial animation through 3D model-based tracking. *ICCV Vision & Graphics*, *1*, 143-150.
- PITTENGER, J. B. (1991). On the difficulty of averaging faces: Comments on Langlois and Roggman. *Psychological Science*, *2*, 351-353.
- RUSSELL, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, *115*, 102-141.
- SPENCER-SMITH, J., INNES-KER, A. H., & TOWNSEND, J. T. (2000, November). *Dynamics of expression*. Paper presented at Object Perception and Memory 2000, New Orleans.
- 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.

NOTES

1. Three-way MDS treats each participant's dissimilarity matrix separately, allowing for large observer differences.
2. Stress is a measure of how well the derived distances fit the distances predicted by the original proximity data.

(Manuscript received November 16, 2000;
revision accepted for publication March 21, 2001.)