COGNITION AND EMOTION, 2004, *18* (5), 701–710

BRIEF REPORT

The dynamic aspects of emotional facial expressions

Wataru Sato and Sakiko Yoshikawa
Kyoto University, Japan

In this study, we presented computer-morphing animations of the facial expressions of six emotions to 43 subjects and asked them to evaluate the naturalness of the rate of change of each expression. The results showed that the naturalness of the expressions depended on the velocity of change, and the patterns for the four velocities differed with the emotions. Principal component analysis of the data extracted the structures that underlie the evaluation of dynamic facial expressions, which differed from previously reported structures for static expressions in some aspects. These results suggest that the representations of facial expressions include not only static but also dynamic properties.

Dynamic changes in the facial expression of emotions are a particularly valuable source of information in face-to-face interactions. They indicate moment-to-moment changes in the emotional state of other individuals.

Based on close observations of thousands of facial expressions, Ekman and his colleagues described various expressions from a dynamic perspective (Ekman & Friesen, 1975; Matsumoto, Ekman, & Fridlund, 1991). For example, Ekman & Friesen (1975) described the changes in the eyes in surprise as follows: ''The eyes are opened wide during surprise . . . When the upper eyelid is raised, exposing the sclera . . . it is almost always a very brief action lasting a fraction of a second'' (p. 40).

Such vivid descriptions suggest that dynamic information plays an important role in the recognition of specific emotions in facial expressions. Consistently, Tomkins (1982) has pointed out that the temporal patterns of physical responses, including facial expressions, are useful cues for distinguishing emotions.

Correspondence should be addressed to: Wataru Sato, Department of Cognitive Psychology in Education, Graduate School of Education, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto, 606-8501, Japan; e-mail: L50158@sakura.kudpc.kyoto-u.ac.jp

We sincerely thank Professor Craig A. Smith and our anonymous reviewers for their helpful advice. This study was supported by Special Coordination Funds for promoting Science and Technology by the Science and Technology Agency of the Japanese Government.

http://www.tandf.co.uk/journals/pp/02699931.html DOI:10.1080/02699930341000176

However, our understanding of such dynamic information in facial expressions, especially in depicting the expression of particular emotions, remains very limited. Although some experimental studies have addressed the temporal characteristics of facial expressions, these studies have not focused on the differences between emotions, but have examined other aspects, such as the differences between spontaneous and deliberate expressions (Hess & Kleck, 1990, 1997; Bugental, 1986; Weiss, Blum, & Gleberman, 1987) or the differences between dynamic and static expressions (Bassili, 1978, 1979; Humphreys, Donnelly, & Riddoch, 1993).

The present study explored the dynamic aspects of facial expressions of emotion from the perceivers' viewpoint. Does the rate at which the face changes affect the recognition of facial expression? Do the temporal characteristics of facial expressions differ among emotional categories? For this purpose, we presented computer-morphed animations of facial expressions to subjects, and asked them to rate the naturalness of the rate of change of each animation clip for that particular emotion. The stimuli consisted of smoothly changing facial images, in which we rapidly presented intermediate frames that seamlessly transformed a face from a neutral expression into one of six emotional expressions. By adopting this method, we were able to systematically manipulate the rate of change. Additionally, this tool has several merits over other methods, such as recorded video clips. For example, because the stimuli were made from static images, they contained little extraneous noise, such as head movement or eye aversion. Since this tool used static images chosen from the Pictures of Facial Affect (Ekman & Friesen, 1976), which is frequently used in studies of emotion, our results can be compared with a wealth of previous findings.

METHOD

Subjects

A total of 43 volunteers (28 women and 15 men; mean age, 20.9) participated in this experiment. All the subjects were right-handed, and had normal or corrected-to-normal visual acuity.

Stimuli

The materials consisted of seven expressions (six emotional and one neutral) on a female (C) and male (JJ) model chosen from the Pictures of Facial Affect (Ekman & Friesen, 1976).

Using these photos as raw materials, we made animated clips of emotional facial expressions. First, we created 49 intermediate images between the neutral (0%) and emotional (100%) expressions, in 2% steps (Figure 1a), using computer-morphing techniques (Mukaida et al., 2000) implemented on a Linux computer. During this procedure, spatial normalisation (i.e., the correction of head position and face size) was used to minimise artifacts. Next, to create a moving clip, we presented a total of 51 images in succession: one neutral image, 49 intermediate images, and the final emotion image.

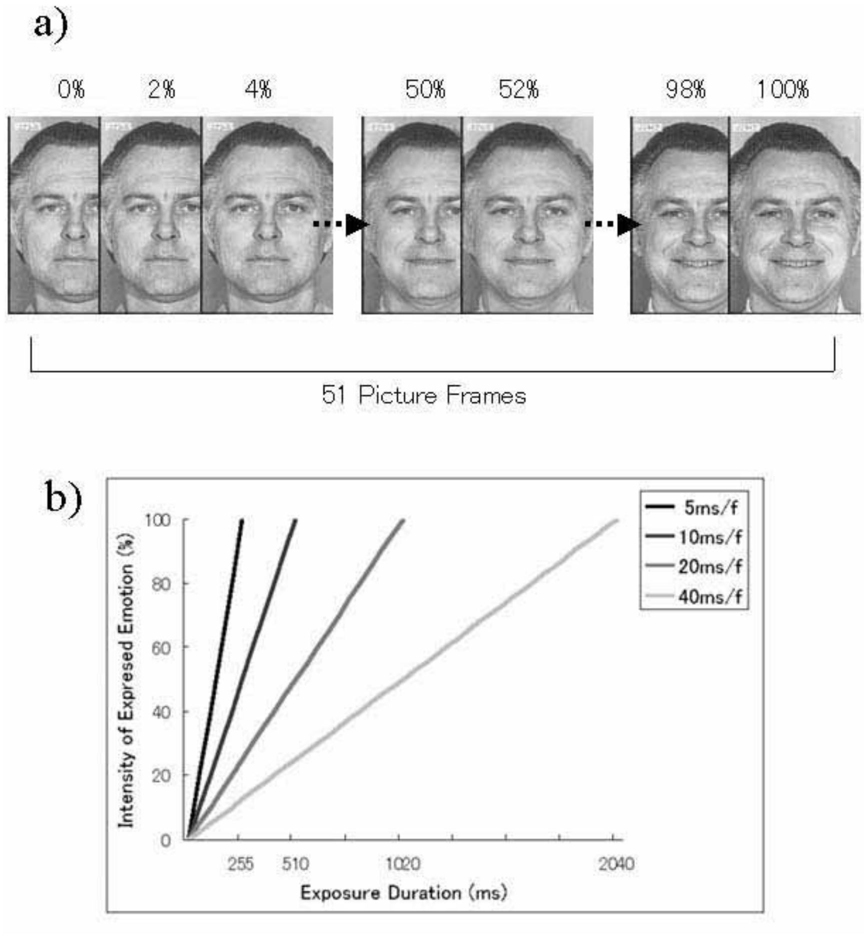

Figure 1. (a) Illustration of a stimulus clip. A total of 51 images that were seamlessly transformed from a neutral expression to one of six emotional expressions were presented. (b) Schematic illustration of the four velocity conditions: 5 ms/frame (total 255 ms), 10 ms/frame (or 510 ms), 20 ms/frame (or 1020 ms), and 40 ms/frame (or 2040 ms). The horizontal axis indicates the stimulus duration and the vertical axis indicates the intensity of the emotion. An intensity of 100% refers to the intensity of the original images from Ekman and Friesen (1976).

Each stimulus was presented at one of four velocities (Figure 1b): 5, 10, 20, and 40 ms/frame, and accordingly lasted 255, 510, 1020, and 2040 ms, respectively.[1]

Apparatus

The events were controlled by a program written in Visual C++5.0, implemented on a Windows computer. The stimuli were presented on a 17-inch flat CRT monitor (600 $\times$ 800 pixels, 256 colour) from a viewing distance of 0.57 m, and subtended a visual angle of 10.5 $\times$ 7.0° (300 $\times$ 200 pixels).

Procedure

The order of presentation of the velocity conditions was counterbalanced across emotions and subjects. The interval between each clip was 1500 ms.

After the researcher had stated the emotion that was to be shown, a sequence consisting of four animated clips was presented. The subjects were tested individually, and were instructed to evaluate each clip in terms of the naturalness of the rate of change of a particular emotion. Each clip was rated on a 7-point scale from 1 (not at all natural) to 7 (very natural). The subjects were allowed to view the sequence repeatedly by clicking a button on the CRT monitor, until they were confident of their rating.

Each subject completed 12 trials corresponding to six basic emotions in each model, and evaluated a total of 48 expression clips. Before the trials, the subjects were given a few practice trials to familiarise them with the procedure. The order in which the emotions were presented was counterbalanced across subjects. The entire series of tasks took about 20 minutes.

RESULTS

The naturalness ratings were subjected to a two-way analysis of variance (ANOVA) (Figure 2a), with emotion (anger, disgust, fear, happiness, sadness, and surprise) and presentation velocity (5, 10, 20, and 40 ms/frame) as within-subject factors. In cases where the assumption of sphericity was not met ($p < .1$, Mauchley's sphericity test), the Greenhouse-Geisser adjusted degree of freedom was used (cf. Greenhouse & Geisser, 1959). The results revealed significant main effects of emotion, $F(5, 210) = 5.91$, $p < .001$, and presentational velocity, $F(1.97, 82.77) = 31.89$, $p < .001$, and a significant interaction of emotion $\times$ presentation velocity, $F(9.17, 385.07) = 11.22$, $p < .001$. For the interaction, follow-up analyses of the simple main effects revealed that the simple main effect of emotion was significant for all presentation velocities, $Fs(5, 840) = 14.90$, 3.66, 8.72, and 9.54 for 5, 10, 20, and 40 ms/frame, respectively; all $ps < .005$, and the simple main effect of presentation velocity was significant for all emotions, $Fs(3, 756) = 17.00$,

[1] These velocities were based on our preliminary studies. Although some previous studies that analysed facial changes suggested that some emotions are expressed within several hundred milliseconds (e.g., Richardson, Bowers, Bauer, Heilman, & Leonard, 2000), details of these temporal characteristics have not yet been examined systematically. It is possible that our manipulation of the stimulus presentation confounds the effect of velicity with that of presentation duration. However, the duration of stimulus presentation has been shown to have less of an impact on the evaluation of emotions in facial expressions (Wild, Erb, & Bartels, 2001).

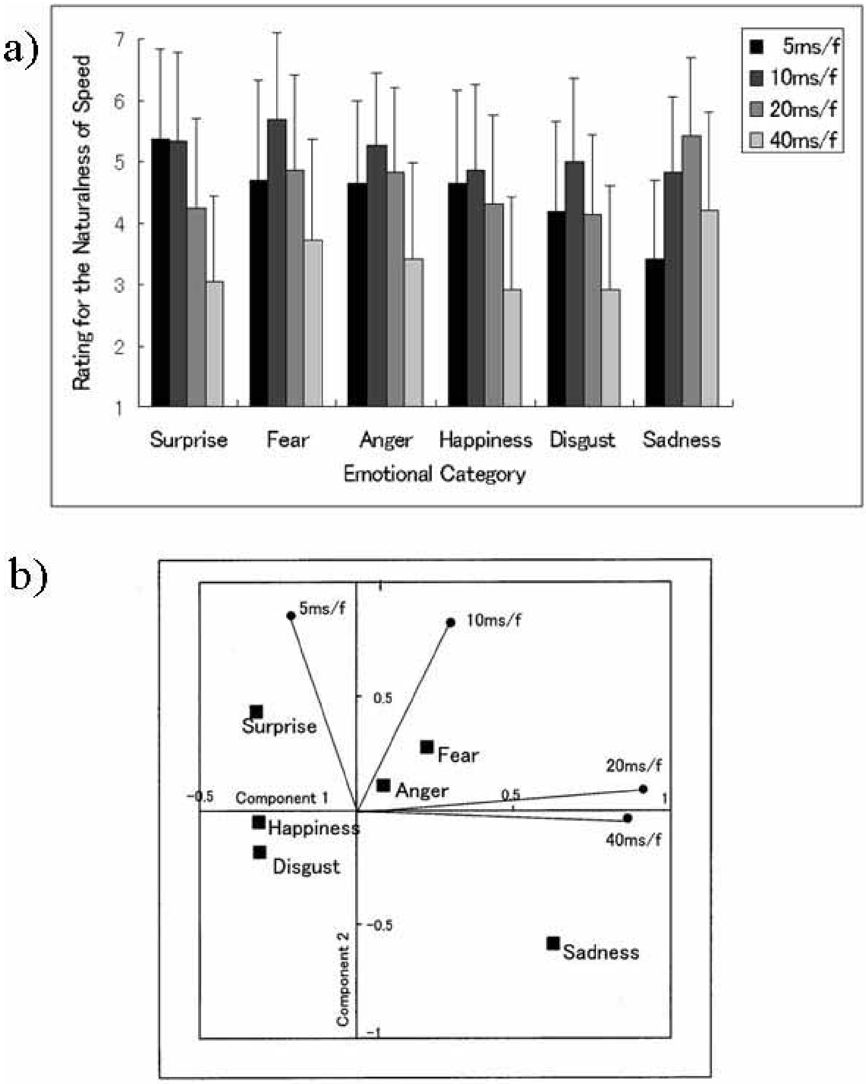

Figure 2. (a) Mean naturalness ratings of the velocity for each expression. The emotions are in ascending order for the values under the fastest condition (i.e., 5 ms/frame). (b) Scatter plots of the mean component scores for each emotion with the plots of component loadings for each velocity condition analysed using the principal component analyses.

19.60, 17.18, 20.69, 19.71, and 32.54 for surprise, fear, anger, happiness, disgust, and sadness, respectively; all ps $< .001$.

For the simple main effect of emotion, *post-hoc* multiple comparisons using Ryan's method ($p < .05$) showed the following. At 5 ms/frame, surprise was rated higher (more natural) than the other five emotions, while sadness was rated lower (less natural) than the other five. At 10 ms/frame, fear was rated higher than happiness, disgust, and sadness. At 20 ms/frame, sadness was rated higher than surprise, happiness, and disgust, and fear and anger were rated higher than disgust. At 40 ms/frame, sadness was rated higher than all the other emotions except fear, and fear was rated higher than surprise, happiness, and disgust.

For the simple main effect of presentation velocity, *post-hoc* trend analyses were used to investigate the profiles of the changes in the ratings across velocity levels. The results revealed strong linear trends for naturalness rating as a function of velocity, for surprise, fear, anger, happiness, and disgust, Fs$(3, 840) = 41.89$, 13.68, 16.16, 25.21, and 18.86, respectively; all ps $< .001$. For these emotions, there were also some significant higher order trends, but their F-values were less than half those of the linear trends. For sadness, a quadratic trend gave the best fit $F(3, 840) = 21.91$, $p < .001$, while the linear trend did not reach significance ($p > .1$). To sum up, sadness was the only exception to the rule that there was a linear decrease in naturalness with decreasing velocity.

To extract the latent structure underlying expressions of emotion, principal component analysis (PCA) with varimax rotation was conducted on the correlation matrix of the naturalness ratings. The results revealed two primary components (Figure 2b) that accounted for 78.20% of the total variance in the data (43.88 and 34.31%, respectively). Based on the plots of component loadings for velocity conditions, these components could be interpreted as the dimension reflecting two slower and faster velocity conditions, respectively. The first component, which appeared to reflect the tolerance for slower velocity, had positive loadings for sadness, fear, and anger, and negative loadings for surprise, happiness, and disgust. The second component, which was interpretable as the appropriateness for faster velocity, had positive loadings for surprise, fear, and anger, and negative loadings for happiness, disgust, and sadness. Additionally, it was interesting to note that the second component corresponded roughly to the previous ''arousal'' dimension in a two-dimension model of facial expressions (Russell, 1997).

DISCUSSION

The main results of the ANOVA and *post-hoc* analyses were as follows: (1) The naturalness rating for each presentation velocity differed across the emotions examined. This was particularly notable at the two extremes: at the fastest speed, surprise was considered the most natural, while sadness was the least natural, whereas at the slowest speed, sadness and fear were judged as more natural than the other emotions. (2) The changing pattern of naturalness with presentation velocity differed among the emotions. Only for sadness were faster presentations not evaluated as a more natural expression of that emotion.

The conspicuousness of surprise, as the fastest expression, and sadness, as the slowest, is in line with the results of a previous observational study (Ekman & Friesen, 1975). Additionally, the speed of surprise concurs with a recent neurophysiological model, which proposes that the startle response is the fastest emotional response and that it is implemented at the subcortical level (Halgren & Marinkovic, 1995). The connection between a slow rate and the perception of sadness are consistent with findings in the literature on vocal emotion recognition (Breitenstein, Lancker, & Daum, 2001).

In contrast to surprise, the high naturalness rating of fear under the slowest condition was an intriguing finding. Although subjects tend to confuse fear and surprise in experiments using static facial images (e.g., Tomkins & McCarter, 1964), Lemay, Kirouac, and Lacouture (1995) found that the confusion between these two expressions was reduced in video presentations of facial images. Ekman and Friesen (1975) also described the temporal difference between fear and surprise as follows: ''the way in which fear differs from surprise is in its duration. Surprise is the briefest emotion, unfortunately fear is not. Surprise has always short duration'' (p. 49). Taken together, the difference between these two expressions may depend primarily on dynamic information.

Moreover, the results of the PCA uncovered the organisation of the representation of dynamic facial expressions. It is intriguing that the second component corresponds roughly to the previous ''arousal'' dimension in a two-dimensional model of facial expressions (Russell, 1997). The two dimensions, pleasure and arousal, were proposed to represent not specific emotions but general features common to different emotions, and to be automatically recognisable in facial patterns (Russell, 1997). However, the meaning of this second dimension varies across studies (cf. Ekman, Friesen, & Ellsworth, 1982), although some explanations of the arousal dimension have been proposed, for example, activity (Russell, 1997) and expressed energy (Adolphs, Russell, & Tranel, 1999). Our results suggest that this dimension is related to the dynamic aspect of emotional faces.

Aside from this similarity between the second component and the arousal dimension, the overall configurations of emotions were rather different from the previous structures of emotional facial expressions. This difference may reflect differences in the stage of processing. Neuropsychological studies (e.g., Peper & Irle, 1997) have shown that emotional expressions are processed in several stages, from the visual analysis of facial features to the activation of emotional representations. Since the previous two-dimensional structures of emotional facial expressions have shown a correspondence with those of emotional words (Russell, 1997), those structures may reflect a structure of emotional meaning that is independent from the visual representations. In contrast, we speculate that our structure describes the organisation of the visual representations of facial expressions, particularly those that reflect visual motion information, because our experimental task had little to do with the emotional content.

As for the visual representations of facial expressions of emotion, the neuroscience literature gives us some suggestions. Monkey single-unit recordings have indicated that particular populations of cells in the superior temporal sulcus (STS) selectively respond to facial expressions of emotion (Perrett & Mistlin, 1990). Similarly, recent functional imaging studies have shown that the homologous human area also responds to both static and dynamic facial images (Haxby, Hoffman, & Gobbini, 2000). It is interesting that this STS region is the site where visual form and motion information are integrated (Oram & Perrett, 1996). This functional anatomy suggests that our neural mechanism represents emotional facial expressions as a combination of dynamic and static visual information.

In our experiment, some stimuli were evaluated as less natural for a particular emotion. When the facial expressions appeared less natural, what meaning did the faces express? Although our results do not speak conclusively to this question, some subjects provided interesting reports during debriefing. They noted that some expressions were inappropriate as an expression of the stated emotion, but were appropriate for expressing another emotion. For instance, some subjects reported that the fastest sad expression did not appear to be an expression of sadness, but rather of perplexity. We find these changes in how the emotions were perceived intriguing and are now planning further studies to investigate this phenomenon.

A particular feature of our methodology was the use of computer-morphed animation. By adopting this tool, we were able to present smooth changes in the facial images. Certainly, our informal survey at debriefing showed that although the subjects viewed the images at a systematically controlled velocity, they did not notice that the images were either artificial or computer-generated. Additionally, as mentioned above, the present method has merits such as the easily adjustable rate of change and lack of noise. At the same time, however, this tool has some limitations. First, all the features on the faces in our stimuli changed at the same time, but we do not know whether real faces actually change in this manner. Second, we used a linear increment across the frames, but we do not know whether faces change in this manner either. At present, quantitative data on the exact temporal function of facial expressions that would answer these questions are lacking. Despite these limitations, our results should encourage the use of computer animation in this area. Additionally, some new technology, such as three-dimensional morphing (Zhang & Cohen, 2000), might remedy the technical problems that we have mentioned.

Although our results and previous observations (Ekman & Friesen, 1975) consistently indicate the importance of the speed of changes in facial expressions, this does not mean that this parameter is the sole temporal characteristic related to the recognition of specific emotions. Previous studies have shown that some temporal parameters, such as the changing rate, apex period, and offset time of facial expressions, could help to differentiate spontaneous and deliberate expressions (Hess & Kleck, 1990, 1997; Weiss et al., 1987; Bugental, 1986). Future studies should investigate the roles of these additional temporal parameters in the recognition of facials expressions of emotion.

There is no doubt that real-life facial expressions are dynamic. However, most of the previous psychological literature has used static facial stimuli in the form of still photographs, although some researchers questioned their ecological validity (e.g., Wallbott & Scherer, 1986). Further study of the dynamic aspects of facial expressions will open a new stage in understanding of how emotional expressions are recognised.

Manuscript received 5 March 2001
Revised manuscript received 22 November 2002

REFERENCES

Adolphs, R., Russell, J. A., & Tranel, D. (1999). A role for the human amygdala in recognizing emotional arousal from unpleasant stimuli. *Psychological Science*, *10*, 167–171.

Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 373–379.

Bassili, J. N. (1979). Emotion recognition: the role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, *37*, 2049–2058.

Breitenstein, C., Van Lancker, D., & Daum, I. (2001). The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cognition and Emotion*, *15*, 57–79.

Bugental, D. B. (1986). Unmasking the ''polite smile'': Situational and personal determinants of managed affect in adult-child interaction. *Personality and Social Psychology Bulletin*, *12*, 7–16.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues.* Englewood Cliffs, NJ: Prentice-Hall.

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

Ekman, P., Friesen, W. V., & Ellsworth, P. (1982). What emotion categories or dimensions can observers judge from facial behavior? In P. Ekman (Ed.), *Emotion in the human face*. (pp. 39–55). New York: Cambridge University Press.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

Halgren, E., & Marinkovic, K. (1995). Neurophysiological networks integrating human emotions. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1137–1151). Cambridge, MA: MIT Press.

Haxby, J. V., Hoffman, E. A., & Gabbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science*, *4*, 223–233.

Hess, U., & Kleck, R. E. (1990). Differentiating emotion elicited and deliberate emotional facial expressions. *European Journal of Social Psychology*, *20*, 369–385.

Hess, U., & Kleck, R. E. (1997). Differentiating emotion elicited and deliberate emotional facial expressions. In P. Ekman & E. L. Rosenberg (Eds.), *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)* (pp. 271–288). New York: Oxford University Press.

Humphreys, G. W., Donnelly, N., & Riddoch, M. J. (1993). Expression is computed separately from facial identity, and it is computed separately for moving and static faces: Neuropsychological evidence. *Neuropsychologia*, *31*, 173–181.

Lemay, G., Kirouac, G., & Lacouture, Y. (1995). Expressions faciales émotionnelles spontanées dynamiques et statiques: comparaison d'études de jugement catégoriel et dimensionnel. [Dynamic and static spontaneous facial expressions of emotion: A comparison of categorical and dimensional judgements]. *Canadian Journal of Behavioural Science*, *27*, 125–139.

Matsumoto, D., Ekman, P., & Fridlund, A. (1991). Analyzing nonverbal behavior. In P. W. Dowrick (Ed.), *Practical guide to using video in the behavioral sciences* (pp. 153–165). New York: Wiley.

Mukaida, S., Kamachi, M., Kato, T., Oda, M., Yoshikawa, S., & Akamatsu, S. (2000). *Foolproof utilities for facial image manipulation* [unpublished computer software]. Kyoto: ATR.

Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology*, *76*, 109–126.

Peper, M., & Irle, E. (1997). The decoding of emotional concepts in patients with focal cerebral lesions. *Brain and Cognition*, *34*, 360–387.

Perrett, D. I., & Mistlin, A. J. (1990). Perception of facial characteristics by monkeys. In W. C. Stebbins & M. A. Berkley (Eds.), *Comparative perception: Vol. 2. Complex signals* (pp. 187–215). New York: Wiley.

Richardson, C. R., Bowers, D., Bauer, R. M., Heilman, K. M., & Leonard, C. M. (2000). Digitizing the moving face during dynamic displays of emotion. *Neuropsychologia*, *38*, 1028–1039.

Russell, J. A. (1997). Reading emotions from and into faces: Resurrecting a dimensional-contextual perspective. In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression. Studies in emotion and social interaction. 2nd series* (pp. 295–320). New York: Cambridge University Press.

Tomkins, S. S. (1982). Affect theory. In P. Ekman (Ed.), *Emotion in the human face* (pp. 353–395). New York: Cambridge University Press.

Tomkins, S. S., & McCarter, R. (1964). What and where are the primary affects? Some evidence for a theory. *Perceptual and Motor Skills*, *18*, 119–158.

Wallbott, H. G., & Scherer, K. R. (1986). Cues and channels in emotion recognition. *Journal of Personality and Social Psychology*, *51*, 690–699.

Weiss, F., Blum, G. S., & Gleberman, L. (1987). Anatomically based measurement of facial expressions in simulated versus hypnotically induced affect. *Motivation and Emotion*. *11*, 67–81.

Wild, B., Erb, M., & Bartels, M. (2001). Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: quality, quantity, time course and gender differences. *Psychiatry Research*, *102*, 109–124.

Zhang, C., & Cohen, F.S. (2000, March). *Face shape extraction and recognition using 3D morphing and distance mapping*. Paper presented at Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000, Grenoble, France.

APPENDIX

The photographs used in this study were chosen from the set in Ekman and Friesen (1976). The ID numbers (and emotional categories) of the stimuli were as follows: C-2-18 and JJ-4-07 (happiness), C-1-18 and JJ-5-05 (sadness), C-1-23 and JJ-5-13 (fear), C-2-12 and JJ-3-12 (anger), C-1-10 and JJ-4-13 (surprise), C-1-04 and JJ-3-20 (disgust), and C-2-03 and JJ-3-04 (neutral).