

Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion

Paul Ekman and Wallace V. Friesen
University of California, San Francisco

Maureen O'Sullivan
University of San Francisco

Irene Diacoyanni-Tarlatzis
Aristotlian University, Thessaloniki, Greece

Rainer Krause
University des Saarlandes, Saarbrücken,
Federal Republic of Germany

Tom Pitcairn
University of Edinburgh, Scotland

Klaus Scherer
University of Geneva, Switzerland

Anthony Chan
Chinese University of Hong Kong

Karl Heider
University of South Carolina

William Ayhan LeCompte
University of Maryland

Pio E. Ricci-Bitti
University of Bologna, Italy

Masatoshi Tomita
Waseda University, Tokyo, Japan

Athanase Tzavaras
Aristotelian University, Thessaloniki, Greece

We present here new evidence of cross-cultural agreement in the judgment of facial expression. Subjects in 10 cultures performed a more complex judgment task than has been used in previous cross-cultural studies. Instead of limiting the subjects to selecting only one emotion term for each expression, this task allowed them to indicate that multiple emotions were evident and the intensity of each emotion. Agreement was very high across cultures about which emotion was the most intense. The 10 cultures also agreed about the second most intense emotion signaled by an expression and about the relative intensity among expressions of the same emotion. However, cultural differences were found in judgments of the absolute level of emotional intensity.

In the last 10 years, opinion has shifted about whether facial expressions of emotion are universal. The earlier view that what a facial expression signifies is completely different from culture to culture (Birdwhistell, 1970; LaBarre, 1947; Leach, 1972) is no longer accepted within psychology, although it is still maintained by some anthropologists (Howell, 1985). Those who have become persuaded by the evidence of universal facial expressions of emotion can cite consistent findings across three quite different types of research. Those who remain skeptical, however, can cite flaws in each. Our study was designed to remedy some of these flaws. We will consider the strengths and weaknesses in each type of research on the universality of facial expressions of emotion.

In one type of investigation (Ekman & Friesen, 1971), mem-

bers of one culture were asked to show how their face would look if they were the person in each of a number of different emotional contexts (e.g., "you feel sad because your child died," "you are angry and about to fight"). Universality was demonstrated when observers in another culture did far better than chance in identifying which emotional contexts the expressions were intended to portray. This finding had unusual import because the persons displaying the expressions were members of a visually isolated New Guinea culture (the South Fore). The ability of Americans to understand these New Guinean expressions could not be attributed to prior contact between these groups or to both having learned their expressions from mass media models.

Three problems limit these findings. First, there has been only one such study. It has not been repeated in another preliterate, visually isolated culture, nor for that matter in a literate, non-Western or Western culture. Second, not all six emotions portrayed were accurately recognized. Anger, disgust, happiness, and sadness were distinguished from each other and from fear and surprise, but the American observers could not distinguish the New Guineans portrayals of fear and surprise. Third,

Paul Ekman's work is supported by Research Scientist Award MH 06092 from the National Institute of Mental Health.

Correspondence concerning this article should be addressed to Paul Ekman, Human Interaction Laboratory, University of California School of Medicine, 401 Parnassus Avenue, San Francisco, California 94143.

the facial expressions were posed, and Mead (1975) argued that establishing that posed expressions are universal need not imply that spontaneous facial expressions of emotion are universal. The next type of research design answered this criticism.

Facial expressions shown by Japanese and by Americans while they watched stress-inducing films (bodily mutilation) and neutral films (nature scenes) were measured. When the subjects in each culture watched the films alone, unaware of a hidden camera, virtually the same facial responses were emitted regardless of culture (Ekman, 1972). However, when a scientist was present when they watched the films, the Japanese more than the Americans masked negative expressions with smiles (Friesen, 1972). In addition to examining spontaneous facial expressions, this study was the first to show how cultural differences in the management of facial expressions (what Ekman & Friesen, 1969, had termed *display rules*) can mask universal facial expressions.

Two problems limit these findings. First, again it is but a single study; no one has yet attempted to replicate it. Also, the mutilation films elicited only a few emotions (disgust and fear), not allowing determination of whether the full range of spontaneous emotional expressions is universal. The next type of research met these two criticisms.

Photographs of facial expressions were shown to observers who were asked to judge the emotion displayed. Very high agreement was found across 12 literate cultures in the specific emotions attributed to facial expressions. The strength of this evidence is its many replications. Unlike the first two kinds of research, this type of study has been repeated in many cultures, by different researchers (Ekman, Sorenson, & Friesen, 1969; Izard, 1971), and with different photographs of facial expression.

Four questions can be raised about the value of such judgment studies in which the same set of faces is shown to observers in different cultures. First, the observers were shown posed rather than spontaneous expressions. This criticism is at least partially met by the fact that universality was also found in one other judgment study (Ekman, 1972) in which the observers saw spontaneous facial behavior. The expressions of the Japanese and American subjects in the study described earlier, in which subjects had watched body mutilation and neutral films, were shown to Japanese and American observers. These observers were asked to judge whether each person's expressions occurred in reaction to a stressful or a neutral film. The judgments made by the Japanese and American observers were highly correlated and did not differ as a function of whether they were interpreting the expressions of their own or the other culture.

Second, all the cultures had some contact either with each other or with media presentations of facial expressions, and therefore their similar judgments could be the consequence of having learned a common set of facial expressions. This criticism is met by judgment studies in two different, visually isolated, preliterate New Guinean cultures, the South Fore and the Dani. The New Guineans discriminated most but not all of the emotions distinguished by the literate-culture observers. In both New Guinean cultures, happiness, sadness, disgust, and surprise were discriminated from each other and from anger and fear. In the South Fore, fear was not distinguished from surprise (Ekman & Friesen, 1971), but this discrimination was made by the other New Guinean culture, the Dani (Heider &

Rosch, reported in Ekman, 1973). The Dani did not discriminate anger from disgust, but the South Fore did.

The next two criticisms raise questions due to limitations in the judgment task that the observers used to register their impressions. The third one is that the judgment tasks might have concealed cultural differences in the perception of secondary blended emotions. Many students of emotion have noted that facial expressions may contain more than one message (Ekman & Friesen, 1969; Izard, 1971; Plutchik, 1962; Tomkins, 1963). The two emotions in a blend may be of similar strength, or one emotion may be primary, much more salient than the other secondary emotion. In prior cross-cultural studies, the investigators presumed that the expressions they showed displayed a single emotion rather than a blend and therefore did not provide those who observed the expressions the opportunity to choose more than one emotion for each expression. Without such data, however, it is not possible to ascertain whether an expression conveys a single emotion or a blend, and if there is blend, whether cultures agree in their judgment of the secondary emotion. Prior evidence of cross-cultural agreement in the judgment of expressions might be limited just to the primary message, not to the secondary blended emotions.

Fourth, despite agreement about which emotion is depicted, there might be differences in the strength of the perceived emotion. Only one cross-cultural judgment study (Ekman, 1972) obtained intensity judgments, and no differences were found. Further investigation is warranted as only five cultures were examined.

To summarize, there has been no cross-cultural study of whether cultures differ in the perception of secondary blended emotions. To do so requires that the observers be allowed to indicate that an expression shows multiple emotions. Our study was designed to fill this gap and also to replicate the finding that the intensity ascribed to an emotional expression is also universal. Observers were asked to judge the emotions shown in each photograph twice, once restricted to a single choice for each expression and once allowed to register up to seven emotions and the relative strength of each.

Hypothesis 1: There will be agreement across cultures about which emotion is shown in each expression when observers are limited to a single choice. This hypothesis simply predicts that earlier findings with a single-choice judgment task will be replicated.

Hypothesis 2: There will be agreement across cultures about which emotion is the strongest one shown in each expression when observers are allowed to register the presence of up to seven emotions. This hypothesis predicts that allowing multiple-emotion judgments will not eliminate cross-cultural agreement.

Hypothesis 3: There will be agreement across cultures about which emotion is perceived as the second strongest emotion in each expression. This prediction is more tenuous, for whether or not there will be universality about the secondary emotion is not implied by the prior evidence. We make this prediction extrapolating from Ekman and Friesen's (1975) finding on Americans that the muscular display in the expression predicted the secondary emotion that was attributed to the expression.

Hypothesis 4: There will be agreement across cultures in the judgment of the strength of an emotional expression. Whereas Ekman and Friesen (1969) described how cultural differences in display rules could lead to differences in the judgment of emotional intensity, Hypothesis 4 is based on their finding (Ekman, 1972) of cross-cultural agreement in intensity judgments.

Method

Facial Expressions Judged

The facial expressions shown to the observers were drawn from three sources: posed emotions, spontaneous expressions, and photographs in which models followed instructions about which muscles to contract. A large pool of photographs were scored with Ekman and Friesen's (1978) Facial Action Coding System to determine the muscular actions that produced each expression. Three pictures were selected for each of six emotions: anger, disgust, fear, happiness, sadness, and surprise. The pictures selected were the best examples of Ekman and Friesen's description of the muscular configurations that universally signal those emotions. The 18 photographs were black-and-white, head-on views showing only the face of Caucasian adult men (5) and women (4) between the ages of 30 and 40. Three of the women contributed expressions to two emotions, and the fourth woman provided one expression. Two of the men contributed expressions for three emotions, one provided expressions for two emotions, and the other two men contributed one expression each.

Observers

Ekman and Friesen sought to include a broad range of diverse cultures. Although the selection finally depended on opportunities where interested scientists volunteered to participate in the study, the 10 countries in which the study was conducted did include eight languages and both Western and non-Western countries. Two of these—Japan and Sumatra (the Minangkabau)—are known (Ekman, 1972; Heider, 1984) to differ considerably from Western cultures in their attitudes about emotional expression. In every country, the observers were of equivalent age and education (college students). The samples were from the Estonian S.S.R. ($N = 85$), Germany (67), Greece (61), Hong Kong (29), Italy (40), Japan (98), Scotland (42), Sumatra (36), Turkey (64), and the United States (30).

Judgment Tasks and Procedure

In each language, the seven English emotion terms were translated into the native language by one person and then translated back by another to verify accurate translation. With two exceptions, the scientist who made the initial translation and who then collected the data was a member of the culture in which the study was run. The exceptions were the Sumatra data gathered by Karl Heider in the Indonesian language from bilingual Minangkabau in Padang, West Sumatra, and the Turkish data gathered by William LeCompte in the Turkish language from subjects in Ankara.

The seven English emotion terms included a single word for each type of expression shown (anger, disgust, fear, happiness, sadness, and surprise) plus contempt. Although previous studies had either not allowed contempt as a response alternative or combined it with disgust, here it was provided as a separate alternative because of other interest in whether contempt can be distinguished from disgust expressions (Ekman & Friesen, 1986).

The photographs were prepared as 35-mm slides so they could be shown to groups of observers. The same random order of presentation was used in every culture. The first time the observers saw the slides, each picture was shown for 10 s, during which the observers were instructed to check on their answer sheets one of the seven emotion terms to register their judgment of each expression. Before observers saw the expressions a second time, the instructions explained that some expressions might show many emotions at the same or different strength, whereas other expressions might show only one emotion. In their second viewing, observers were instructed to rate each of the seven emotions in terms of whether it was absent or present, and if it was present to indicate its strength on an 8-point scale from *slight* (1) through *moderate*

Table 1

Single-Emotion Judgment Task: Percentage of Subjects Within Each Culture Who Chose the Predicted Emotion

Nation	Happiness	Surprise	Sadness	Fear	Disgust	Anger
Estonia	90	94	86	91	71	67
Germany	93	87	83	86	61	71
Greece	93	91	80	74	77	77
Hong Kong	92	91	91	84	65	73
Italy	97	92	81	82	89	72
Japan	90	94	87	65	60	67
Scotland	98	88	86	86	79	84
Sumatra	69	78	91	70	70	70
Turkey	87	90	76	76	74	79
United States	95	92	92	84	86	81

(4) to *strong* (8). This time the slides were shown for 30 s each, during which the observers made judgments about all seven emotions for each expression.

Results

Replicating the Findings of Universality

Because there were 3 expressions for each of 6 emotions judged by members of 10 cultures, there were 180 opportunities for the cultures to agree with Ekman and Friesen's predictions and with each other about which emotions are universally signaled by each facial expression. Considering first the single-choice judgments, the emotion term chosen by the majority of the subjects in each culture was, as predicted, 172 of 180 times. This high level of agreement across cultures supports Hypothesis 1, replicating previous findings that also used a single-choice judgment task.

Table 1 summarizes the results collapsed across the three expressions for each type of emotion, listing the percentage within each culture who gave the predicted emotional judgments. The figures in Table 1 are within a few points of what was reported 15 years ago with different photographs and observers (Ekman et al., 1969; Izard, 1971). Although there is some variation in the extent of agreement, what is most relevant to Hypothesis 1 is that the majority of the observers in every culture judged the emotions as predicted.

Although these descriptive data very strongly support Hypothesis 1, we also computed kappa coefficients (Hubert, 1977) to obtain a test of significance. Kappa evaluates the extent to which the judgments were as predicted. We prepared 7×7 tables for each culture, plotting for each of the seven emotions the distribution of obtained against predicted judgments and pooling the judgments across the three photographs depicting each emotion. In all 10 cultures, the kappas were significant beyond the .001 significance level. These are shown in the first column of Table 2. To be certain that pooling results across photographs did not conceal disagreements in the judgment of some of the facial expressions intended to signal a particular emotion, kappas were also computed separately for every photograph for all 10 cultures. Of the 180 kappas (18 photographs \times 10 cultures), 178 were significant beyond the .01 level.

Hypothesis 2 predicted the same findings even when observers were allowed to choose more than one emotion, judging the

Table 2
Kappa Coefficients

Nation	Single judgments	Multiple judgments
Estonia	.790	.744
Germany	.736	.739
Greece	.762	.789
Hong Kong	.763	.718
Italy	.800	.783
Japan	.693	.678
Scotland	.815	.809
Sumatra	.657	.541
Turkey	.729	.738
United States	.835	.607

Note. All figures are significant beyond .001.

intensity of every emotion. To test Hypothesis 2, we determined whether the emotion with the most intense rating was the emotion predicted by Ekman and Friesen and was the same across cultures. Hypothesis 2 was supported; in 177 of 180 times, the emotion rated strongest by the largest number of observers in each culture was the predicted emotion. This is the first evidence of cross-cultural agreement about the most intense emotion when observers can choose more than one emotion.

Kappa coefficients were also computed for the judgments made on the intensity scales. For each observer, the score used was the emotion scale rated as strongest. Again, all 10 kappa coefficients were significant. Table 2 lists the kappa coefficients computed on the single judgment data and the multiple judgment data.

Is There Agreement About the Second Most Intense Emotion?

Hypothesis 3, which predicted universality in the secondary emotion, could be tested only with those expressions that the observers had judged as showing more than one emotion (i.e., that were blends rather than single-emotion expressions). Although the selection of photographs had followed Ekman and Friesen's (1978) guidelines for excluding blends in which two emotions are signaled with equal strength, those guidelines were not designed to exclude blends in which a secondary emotion is conveyed with less strength than is the primary emotion. Only judgment data in which the observers were allowed to record the presence of more than one emotion and the relative strength of each emotion can reveal whether an expression conveys a single emotion or a blend.

We set the following criteria for classifying an expression as a blend, pertinent to testing Hypothesis 3: (a) The second strongest emotion had to have a mean of at least 1.5 on the *absent* (0) to *strong* (8) intensity scale; (b) at least half of the judges within a culture had to contribute to that rating; and (c) at least two cultures had to meet the first two criteria. There were 180 opportunities for the judgments (18 expressions \times 10 cultures) to meet these criteria.

Our criteria were met 98 times, involving the judgments of 13 of the 18 expressions. None of the photographs selected to signal happiness met the criteria for signaling a secondary blended emotion. The judgments of the sadness and the surprise

photographs met the criteria for signaling a secondary blended emotion too infrequently for inclusion in the analysis of Hypothesis 3 (for sadness, only 5 of 30 opportunities; for surprise, only 8 of 30). There was complete agreement across the 10 cultures about the secondary emotion signaled by the disgust and by the fear expressions. In every culture on every expression of disgust, the secondary emotion was contempt. In every culture on every expression of fear, the secondary emotion was surprise. Whereas all three anger expressions met the criteria for signaling secondary blended emotions in nearly every culture, the secondary emotion varied with the expression. Disgust was the blended emotion on one anger expression, surprise on another anger expression, and on the third expression the cultures disagreed about the secondary blended emotion (four judged it to be contempt, four judged it to be disgust, and two did not see any secondary emotion).

Cultural Differences in the Intensity of the Judged Emotion

To test Hypothesis 4's prediction of universality in intensity of emotion judgments, we computed a one-way multivariate analysis of variance (MANOVA) with culture as the independent variable and the mean intensity of the six emotions as the six dependent variables. SPSSX MANOVA was used for the analysis. By use of Wilks's criterion, there was an overall effect for culture, $F(54, 2743) = 3.95, p < .001$. The results reflected a moderate association between culture and the intensity judgments of emotion ($\eta^2 = .32$; Tabachnick & Fidell, 1983). Six univariate F tests ($dfs = 9, 542$) ranged from 2.93 ($p < .002$) for sadness to 6.66 ($p < .000$) for surprise, indicating significant differences among cultures for each of the six emotions.

Rather than examining all the cell means on an atheoretical, pair-wise basis, we used post hoc comparisons to address two possible explanations for the significant overall and univariate F s. The first possibility was that people will judge a foreigner's expressions to be less intense than expressions shown by members of their own culture. Attributions of less intense emotions to foreigners might be due to politeness or to greater uncertainty about the emotional state of a person from a culture with which one is less familiar. In our experiment, the three Asian cultures could clearly recognize that the Caucasians in the photographs were not from their own culture. In the first post hoc comparison, therefore, we used Scheffé's procedures to contrast the mean intensity ratings of the three Asian cultures with the mean intensity ratings of the other seven cultures for each of the six emotions. Table 3 shows that the intensity ratings made by the Asian and non-Asian cultures were significantly different for fear, happiness, and surprise. Although the Scheffé test was performed on the separate cell means for each culture, Table 3 gives the average intensity ratings for the two contrasting cultures to clarify the differences between them. Even those that were statistically significant are numerically small. None of the differences were as great as a full point on the 9-point intensity scale.

A second explanation of the significant MANOVA is that observers who made their judgments in languages other than English would give different intensity judgments than would those making their judgments in English. The mean intensity ratings of the English-speaking cultures (Scotland and the United

Table 3
Post Hoc Analyses of Intensity Ratings When Judging Foreigners and Nonforeigners

Emotion	Asian countries	Non-Asian countries	<i>F</i>	<i>p</i>
Happiness			19.63	<.05
<i>M</i>	6.20	6.68		
<i>SD</i>	1.46	1.19		
Surprise			35.25	<.01
<i>M</i>	5.22	6.01		
<i>SD</i>	1.66	1.54		
Sadness			0.61	<i>ns</i>
<i>M</i>	5.69	5.93		
<i>SD</i>	1.63	1.66		
Fear			42.22	<.01
<i>M</i>	5.73	6.70		
<i>SD</i>	2.00	1.47		
Disgust			9.23	<i>ns</i>
<i>M</i>	5.64	6.14		
<i>SD</i>	1.73	1.56		
Anger			5.76	<i>ns</i>
<i>M</i>	5.80	6.04		
<i>SD</i>	1.66	1.61		

Note. Grouped means for the Asian and non-Asian cultures are provided for ease of comparison. In the post hoc analyses, we used Scheffé's procedures to contrast the separate cell means of the three Asian cultures with those from the seven non-Asian cultures. The *F* values were evaluated using $F' = (k - 1) F_c$, with $k = 10$. Degrees of freedom were 9, 542.

States) were compared with those of the other eight countries by using Scheffé's procedures. None of these *F*s (*d*f_s = 9, 542) was significant, suggesting that language differences among cultures is unimportant in judging the intensity of emotional expressions.

Another way to search for cultural differences in intensity judgments was to look for any disagreements about which of two expressions showing the same emotion was the most intense. We set the following criteria for including expressions in this analysis: Two expressions of the same emotion had to be judged as differing in mean intensity (a) by at least one point and (b) in at least two cultures. When that happened, we determined whether the direction of that difference was the same in those two cultures and in all the other cultures. Although we tallied the results across all cultures if our criteria were met, rarely did we find that the mean intensity difference was greater than a point in more than 3 of the 10 cultures. For example, on two of the faces the mean intensity ratings on the anger scale was 6.6 and 7.7 for the Scots and 5.9 and 7.5 for the Italians. We therefore determined whether the face that was rated as most intense was the same for the Scots and the Italians and for the other 8 cultures, even though in these other 8 cultures the ratings of these two pictures did not differ by as much as one full point. Whenever the mean intensity ratings for a pair of expressions depicting a particular emotion differed by less than a point in every culture, the pair of expressions was considered one in which the faces were rated as the same intensity across cultures.

There were 130 opportunities for disagreement about which of a pair of expressions showing the same emotion is the most intense (13 pairs of photographs in which the mean ratings for a pair of expressions differed by at least a point for 2 cultures ×

10 cultures = 130). The 10 cultures agreed about which was the most intense expression 119 of 130 times (binomial test, $z = 9.47$, $p < .0001$).

Discussion

The main, consistent, and robust finding was agreement across cultures in their interpretation of facial expressions of emotion. Three new findings support the view that there are universal facial expressions of emotion. First, cross-cultural agreement is not dependent on limiting observers to choosing only one emotion for each expression. Even when observers were allowed to indicate that an expression showed many emotions, agreement was very high about which emotion was the strongest. Second, cross-cultural agreement is not limited to just the strongest emotion expressed by a face. There was very high agreement across cultures about the second strongest emotion signaled by an expression. Third, cross-cultural agreement is not just about which emotion an expression displays but also about the relative strength of expressions of the same emotion. With few exceptions, the cultures agreed about which of two different expressions of the same emotion was the most intense.

One possible limit on these findings, however, is that all of the observers were college students, all of whom had been exposed to some of the same mass media depictions of facial expressions. Nearly 20 years ago, concerned that their findings of universal facial expressions might be attributed to the opportunity to learn the meaning of expressions from mass media examples rather than as a consequence of evolution, Ekman and Friesen (1971) examined observers in a visually isolated, preliterate culture. They found that judgments of anger, disgust, fear, sadness, and happiness made by these preliterate people were no different than judgments made by college students in eight literate cultures. Given this data base, it is quite unlikely that less educated persons in the cultures we studied would provide different judgments. Nevertheless, data on such observers would make our findings more conclusive.

Further research is needed also to explore alternative explanations of why secondary emotions were perceived for the disgust, fear, and anger expressions but not for the happy expressions and very few for either the sadness or surprise expressions. This difference among emotions might be due to greater similarities among some emotions in appearance or semantic connotations or, less interestingly, to nonreplicable idiosyncracies in the samplings of expressions in this particular study. The secondary emotions found for the disgust and fear expressions were consistent with past studies in which either a single-emotion choice or ratings were obtained (see Ekman, Friesen, & Ellsworth, 1972, chapters 13 and 14, for a review of more than a dozen earlier studies and Russell & Bullock, 1985, for more recent work). However, the failure to find secondary emotions for the expressions of happiness and many secondary emotions for either the sad or the surprise expressions was not consistent with these past studies.

There was also some evidence of cultural differences in intensity judgments. There was some support for the idea that observers attributed less intense emotions to expressions that they could tell were shown by foreigners. The Asians obviously knew that the Caucasians shown in the photographs were members of a foreign culture. Without showing Asian expressions to

Caucasians, we can not know if making less intense emotion attributions to the expressions shown by a foreigner is common across all the cultures studied or unique to the Asian cultures. And without showing Asian faces to Asians, we cannot know if making less intense emotion attributions is general to the Asians judgments of anyone or just of foreigners. In any case, the fact that the Asian and Caucasian ratings did not differ significantly on anger, disgust, or sadness casts some doubt on this line of reasoning. There is no obvious reason that these three emotions would not be influenced by any general tendency to underestimate the strength of emotional expressions shown by a foreigner. The possibility must be considered that the cultural differences in intensity judgments we found are not of real significance despite their statistical significance. Although some differences were statistically significant, even they were very small.

We (Ekman & Friesen, 1969) have always maintained that facial expressions of emotion are both universal (in the evolved muscular displays for each emotion) and culturally variable (in the display rules, some of the antecedents, coping, memories, etc.). The evidence now for universality is overwhelming, whereas that for cultural differences is sparse. Three changes in the research design might help to reveal cultural differences: (a) Each facial expression of emotion should be shown by people who vary in race, sex, and age; (b) facial expressions should be studied in which the muscular signs of the emotion are registered in only one part of the face; and (c) cultures selected for study should be those in which ethnographic investigations have revealed differing attitudes about the experience or expression of specific emotions.

It is possible, however, that despite such changes, the research design itself is not a sensible one for revealing cultural differences. Taking a facial expression out of social context; eliminating the simultaneous speech, vocal clues, and body movements; freezing the expression in a still photograph; forcing attention to it; and asking for judgments by a detached uninvolved observer may remove many of the sources of cultural differences in the interpretation of facial expression. When we sought to demonstrate how cultural differences in display rules produce different facial expressions (Ekman, 1972; Friesen, 1972), we did not ask people to judge photographs but instead observed how facial expressions change in different social contexts. That study may provide a model for the methods needed to reveal further cultural differences in facial expressions of emotion.

References

- Birdwhistell, R. L. (1970). *Kinesics and context*. Philadelphia: University of Pennsylvania Press.
- Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation*, 1971 (Vol. 19, pp. 207-282). Lincoln: University of Nebraska Press.
- Ekman, P. (1973). Cross-cultural studies of emotion. In P. Ekman (Ed.), *Darwin and facial expression: A century of research in review* (pp. 169-222). New York: Academic Press.
- Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1, 49-98.
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17, 124-129.
- 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. (1978). *Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., & Friesen, W. V. (1986). A new pan-cultural expression of emotion. *Motivation and Emotion*, 10, 159-168.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings*. New York: Pergamon Press.
- Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotions. *Science*, 164(3875), 86-88.
- Friesen, W. V. (1972). *Cultural differences in facial expressions in a social situation: An experimental test of the concept of display rules*. Unpublished doctoral dissertation, University of California, San Francisco.
- Heider, K. G. (1984, November). *Emotion: Inner state versus interaction*. Paper presented at the meeting of the American Anthropological Association, Denver, CO.
- Howell, S. (1985, June). *The face as social fact*. Paper presented at British Psychological Society Conference on the Meaning of Faces, Cardiff, Wales.
- Hubert, L. (1977). Kappa revisited. *Psychological Bulletin*, 84, 289-297.
- Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.
- LaBarre, W. (1947). The cultural basis of emotions and gestures. *Journal of Personality*, 16, 49-68.
- Leach, E. (1972). The influence of the cultural context on nonverbal communication in man. In R. Hinde (Ed.), *Nonverbal communication* (pp. 315-344). London: Cambridge University Press.
- Mead, M. (1975). Review of *Darwin and facial expression* P. Ekman (Ed.). *Journal of Communication*, 25(1), 209-213.
- Plutchik, R. (1962). *The emotions: Faces, theories, and a new model*. New York: Random House.
- Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology*, 48, 1290-1298.
- Tabachnick, B. G., & Fidell, L. S. (1983). *Using multivariate statistics*. New York: Harper & Row.
- Tomkins, S. S. (1963). *Affect, imagery, consciousness* (Vol. 2). New York: Springer Publishing.

Received April 4, 1986

Revision received October 9, 1986

Accepted March 13, 1987 ■