SEMIOTICA III
February 1971
1

PAUL EKMAN, WALLACE V. FRIESEN, AND SILVAN S. TOMKINS

Facial Affect Scoring Technique: A First Validity Study[1]

In 1862 Duchenne published his *Mécanisme de la physionomie humaine, ou analyse électro-physiologique de l'expression des passions*, in which he used "the electrical currents for contraction of the muscles of the face to make them speak the language of passions". In this atlas of the anatomy of emotion, Duchenne delineated many of the muscles whose contractions together contributed to the production of each specific emotional expression of the human face. Although Duchenne was the father of modern kinesiology, of whose work Darwin (1872) had written, "no one has more carefully studied the contractions of each separate muscle and the consequent furrows produced on the skin", later investigators virtually ignored his contributions.

Rather than continuing the study of the particular facial muscles which distinguished one from another emotion, the past five decades of psychological research on the face in relation to emotion instead have been devoted primarily to determining what emotion observers can judge from the face (whether their judgments are accurate, what categories of emotion can be judged, etc.). Comparatively few investigators have actually looked at the face itself, applying some measurement procedure

[1] This research was supported by research grants from the National Institute of Mental Health (MH 11976), and the Advanced Research Projects Agency (AF AFO SR-1229-67), and by a Research Scientist Development Award Type II (5-KO2-MH-06092), and a Research Scientist Award (K5-23, 797) from the Research Fellowships Branch of the National Institute of Mental Health. The authors are grateful to Jerry Boucher, Phoebe Ellsworth Diebold, and Virginia Sullwold for their help in the initial steps of the research, and to Patricia Garlan for her editorial assistance.

to the appearance of the face and determining whether the resulting measurements were related to some index of emotion.

Those few who did measure facial components asked one of two questions, depending upon which index of emotion they employed. (1) Can measurements of the face distinguish among emotions experienced, when the emotional state of the person is defined by his self-report or experimentally varied changes in his environment (Fulcher, 1942; Landis, 1924; Landis and Hunt, 1939; Leventhal and Sharp, 1965; Thompson, 1941; Trujillo and Warthin, 1968)? (2) Can facial measurements predict how observers will judge emotion, or be correlated with observers' judgment of emotion (Frijda and Philipszoon, 1963; Frois-Wittman. 1930)?

All but one of these studies that did examine the face itself obtained positive results. The only negative answer was from the first study by Landis, but that experiment has been severely criticized on a number of methodological grounds (Arnold, 1960; Coleman, 1949; Davis, 1934; Ekman, Friesen and Ellsworth, 1971; Frois-Wittmann, 1930; Honkavaara, 1961; Murphy, Murphy, and Newcomb, 1937). Although all the other studies did obtain positive results, a comparison of their findings does not yield much evidence about how specific facial behaviors are related to specific emotions. This lack of convergence may be due to limitations in the design of each of these experiments, which could have led to idiosyncratic findings. All of the studies were limited by insufficient sampling of at least one of the following: the number of emotions studied (whether defined by observer judgment or environmental condition); the number of different faces measured; the number of different stimulus persons used; and the number and kind of facial components measured. This last limitation may be due to the atheoretical bias of these investigators, none of whom specified the basis for their choice of facial components or scores. Instead, it would seem that they measured anything that appeared to vary in their set of faces, or that they could reliably measure, or that was salient to them. If they had looked to theory, they would have found only some guidance, primarily from Duchenne (1862) and those who were influenced by him (Darwin, 1872; Plutchik, 1962; Tomkins, 1962, 1963). However, even these theorists about emotion who emphasized the face failed to describe completely all of the facial configurations they thought relevant to all of the emotions they discussed. And their descriptions were not intended to be a measurement procedure for determining what emotion is shown in a face.

PROBLEM

Our purpose was to develop a tool for the measurement of facial behavior that would be equally applicable to predicting observers' judgments of emotion and to distinguishing emotional state as indicated by environmental condition or self-report. The Facial Affect Scoring Technique (FAST) was to be constructed for use with both still and motion records of the face. A second aim was to provide as complete as possible a description of the facial components that would distinguish among emotions to facilitate further development of theory.

We will report here on the development of FAST and describe a test of its validity in predicting the recognition value of a face. Subsequent reports will deal with other validity questions, such as distinguishing among faces in terms of self-reported or posed emotion, and emotion-arousing circumstances.

FAST's validity in predicting the recognition value of a face was investigated by determining whether FAST scores could distinguish the emotions ascribed to faces by observers. Only faces that had yielded high agreement among observers about the presence of a single emotion were considered. An attempt was made to include faces of as many different stimulus persons as was possible, to evaluate whether the facial scoring technique could cope with physiognomic differences.

Our hypothesis was that FAST scores would accurately predict the judgments of the specific emotions shown on the stimulus persons' faces.

METHOD

The development of the facial affect scoring technique (FAST)

Our first decision was to develop FAST in terms of emotion categories (happiness, anger, surprise, etc.) rather than emotion dimensions (pleasantness-unpleasantness, active-passive, etc.). The decision was based on three considerations. First, more past work specifying facial components has utilized emotion categories than dimensions. Second, recent cross-cultural research has suggested that there are a set of facial components that are associated with emotion categories in the same way for all men, since the same faces were found to be judged as showing the same emotions in many cultures (Ekman, 1968; Ekman, Sorenson, and Friesen, 1969; Ekman and Friesen, 1970; Izard, 1968; 1970a, 1970b). Finally, our own theoretical bias was to conceptualize emotions in terms

of categories rather than dimensions. The specific categories selected were those that had been consistently found by all the investigators within Western cultures who had attempted to determine how many categories of emotion can be judged from the face, and that had also been found by both Ekman *et al.*, and Izard to be interpreted in the same way across cultures. These six categories of emotion are: happiness, sadness, anger, fear, surprise, and disgust. These six categories do not include ALL of the emotions that have been found to be reliably judged across cultures, but they are a sufficient number to meet our aim of showing that it is possible to measure the facial behavior that distinguishes one emotion from another.

The next step required a decision about how to describe facial behavior. Initial attempts to describe facial behavior in terms of muscle movements revealed that it was often quite difficult to determine which muscles had moved by looking at the face. A decision was made to describe the appearance of the face primarily in terms of wrinkles, of tension or relaxation in specific features, and of positions of features. These appearance descriptions were made separately for three facial areas: brows-forehead; eyes-lids-bridge of nose; lower face, consisting of cheek-nose-mouth-chin-jaw. There were two reasons for dividing the face into these three areas, rather than into two or four or five areas. First, this division is based upon the anatomical possibilities for independent movement and appearance changes in the face. Within each of the three areas, a movement in one facial part will almost always lead to an observable change in appearance in other parts of that same area. But across areas, movement in one area (e.g., brow-forehead) need not necessarily result in an observable shift in the appearance of another area (e.g., lower face). While the eye area is most subject to reflecting movements from other facial areas, nevertheless, the eye area can shift in appearance without any observable shift in the other two areas.

The second reason for dividing the face into these three areas was our expectation that if scoring of particular facial components was not to be influenced by impressions of the whole face or from one very pronounced component in one part of the face, it would be necessary to limit the scorer to seeing only one area of the face at a time. Division of the face into three areas appeared a feasible way to mask from view parts of the face when scorers measured still photographs, films, or videotapes.

Lists of facial components within each of the three facial areas for each of the six emotions were compiled. These lists were based in part no past literature, including the theoretical writings and empirical studies

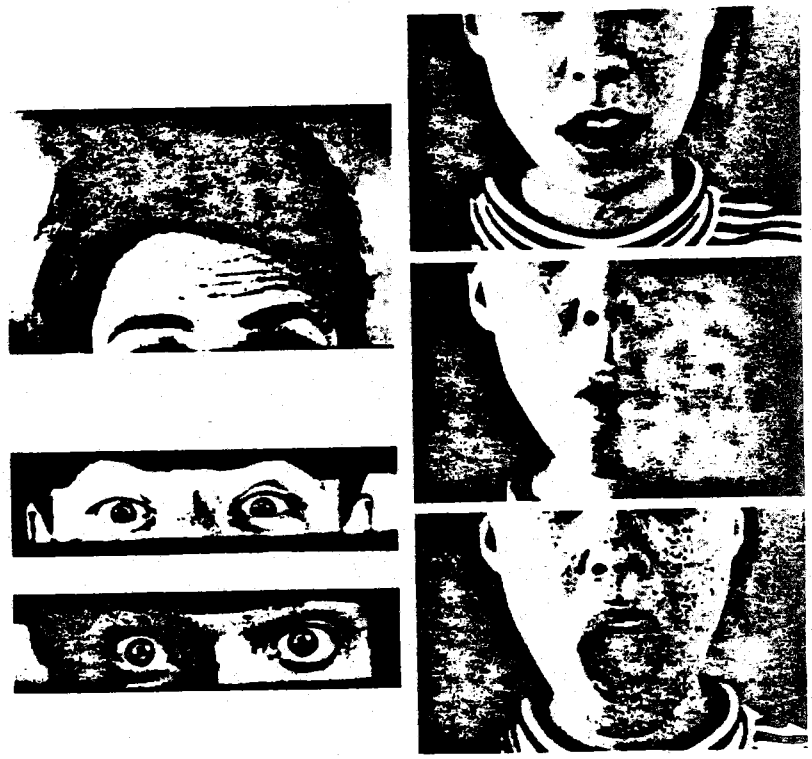

Fig. 1
FAST Items for Surprise

referred to earlier and other writings (Allport, 1924; Birdwhistell, 1952). To this literature, we added our own combined observations and intuitions. We considered for each emotion each of the facial areas and the possibilities for muscular movements within each facial area, checking our hypotheses with a mirror and by looking at each other. An attempt was made to include components for slight, moderate, and extreme intensity versions for HAPPINESS, SURPRISE, and DISGUST, and in the case of ANGER, for both controlled and uncontrolled versions.

Specifying the appearance of some of the components within a facial area required such lengthy, awkward descriptions, which sometimes even then were insufficient, that it became obvious that the only way to make clear to scorers just what was being described was to imitate the appearance ourselves or show a photographic sample. For example, there was no obvious way to describe in words the differences in the appearance of the raised lower eye lids when they were raised by muscular tension or by the skin being pushed up from a grin in the lower face. After a year of pilot studies, in which scorers matched verbal descriptions with photographs, the verbal descriptions were abandoned, and in their place a visual-photographic definition of each item was made.

One or two photographic examples were made to define each item within each facial area. Models were told and shown what to do with the face, rather than asked to pose emotions. A model's attention was focused upon producing a particular appearance in a single facial area. More than one model was employed, since no single person was able to show all of the components in all of the facial areas. (We do not know whether this limit in facial performance might be due to differences in physiognomy, musculature, or learned ability for voluntary control of facial movements.)

FAST consists of three sets of photographs, one set for each facial area, and a few verbal descriptions of head orientation and gaze direction, which, because they were easy to describe verbally, did not require visual definition. Table I shows the number of FAST photographs, or items, in each facial area for each emotion. Figure 1 shows, as examples, the brow area, eye area, and lower-face area items for SURPRISE. Table II gives the head orientation and gaze direction code. The procedure required the scorer to compare part of the face to be scored, one of the three areas, with the set of FAST photograph-items for that facial area, selecting the FAST item which was the best match to the face to be scored. That item then became the score for that facial area. The scoring procedure will be described in more detail below.

TABLE I

Number of FAST Photographic Items for Each Emotion and Facial Area

	Brow-Forehead	Eyes-lids	Lower face
Happiness	1	4	5
Sadness	8	8	10
Surprise	1	2	4
Anger	1	3	9
Disgust	1	2	11
Fear	1	2	4
	13	21	43

TABLE II

FAST: Head/Eye Orientation

Head/Eye Juxtaposition

Up — Down
Chin up and eyes fixated ahead
Chin up and eyes down
Chin down and eyes straight or up
Chin down and eyes down
Chin level and eyes down
Chin up or level and eyes up
None of these

Left — Right
Head oriented left or right with eyes opposed
Not relevant

Head Position Only

Head pulled back
Head thrust forward
Neither of these

Eye Deviations

Wall-eyed
Cross-eyed
Neither of these

Selection of faces to be scored with FAST

Two principles guided selection of the faces to be used to determine whether FAST scores predict how observers judge emotion. Each face should be one which observers agree shows only one emotion, and faces of many different stimulus persons should be included. Other work has shown that faces vary not only in the level of agreement they elicit from observers in judgments of emotion categories, but also in terms of whether they show a single emotion or a blend of two or more emotions (Ekman and Friesen, 1969a, 1969b; Nummenmaa, 1964; Plutchik, 1962; Tomkins and McCarter, 1964). While it would be important to determine whether FAST could differentiate blends from single-emotion faces, and one type of blend from another, this seemed to be a more complex differentiation, which should follow rather than precede determination of whether FAST could distinguish among the emotions conveyed by single-emotion stimuli.

The rationale for using as many different stimulus persons as possible was to test FAST's ability to ignore differences in facial appearance associated with age, sex, and physiognomy, and differences in lighting and contrast that would occur across pictures of different people. Our aim was to obtain 10 pictures of each of the six emotions, using sixty different people. Photographic candidates were selected from the photographs used by past investigators (Ekman and Friesen, 1968; Engen, Levy, and Schlosberg. 1957; Frijda, 1969; Frois-Wittmann, 1930; Izard, 1970a; Nummenmaa, 1964; Tomkins and McCarter, 1964). Any stimulus that had been found to elicit at least seventy percent agreement among observers about the presence of a single emotion was considered a candidate. The norms utilized were published or provided by investigators, including ourselves. The photographic candidates were shown to a new group of judges to replicate the high agreement finding, this time utilizing one judgment task and one group of observers for all stimuli, and to screen the stimuli further with a judgment task more sensitive to discriminating single from blend stimuli.

Pilot studies had revealed that the use of a single choice emotion judgment task could sometimes conceal the presence of a blend. If, for example, a stimulus showed primarily disgust but also signs of anger somewhat less salient, then a single choice judgment task would yield high agreement about only the more salient message, disgust. But, if observers were allowed to indicate a second choice, few would choose disgust only; many would choose disgust as their first choice and anger

as their second choice. Such a two-choice judgment task would, of course, isolate stimuli that showed only one affect.

The candidates selected on the basis of previous norms with a single choice judgment task were shown to a group of 82 observers who were given a two-choice judgment task. The judgment task required them to circle the single word that best described the emotion shown in the face and, IF there was a second emotion also shown, to write a number "2" beside a second word. The emotion words listed in the judgment task were Happiness, Sadness, Anger, Fear, Surprise, Disgust, and Contempt.

Utilizing the results from these observers, and screening stimuli so as to select only single emotion photographs, we could not achieve our goal of 10 pictures for each emotion, each shown by a different stimulus person; further, the standards for inclusion in terms of agreement among observers had to be modified for each of the emotion categories. Table III shows that three criteria were employed to select single emotion stimuli, that the level of agreement varied for the different affect categories and that for the last two categories, Disgust and Fear, it was not possible to obtain 10 pictures, even though the standard for inclusion was considerably more lax for these two categories, allowing the possibility of blend stimuli. (Variations in agreement on a single emotion did have a bearing on the accuracy of FAST predictions, as will be discussed later.) In the 51 pictures selected, there were 28 different persons shown, some more than once within an emotion category or more than once across categories.

Procedure

The photographs were projected life-size by a 2 × 2 slide projector on a back projection screen. The screen was masked so that only one facial area was visible; all pictures were scored on that facial area before the next facial area was presented. Scoring all stimuli on one facial area at a time was intended to reduce the likelihood that a scorer would be influenced in scoring one part of the face by being able to remember how he had scored another part of that same face. To further impair memory for individual pictures, the 51 selected photographs were grouped with 50 other photographs for scoring. These other photographs were often of some of the same stimulus persons. Thus, all 101 pictures were scored on a single facial area before the next area was shown.

TABLE III

Judgment Results on Photographs Selected for FAST Validity Test

Emotion category	Number of stimuli	Number of different persons within each category	Range of percentages saying only the relevant emotion is present	Range of percentages judging the next most frequent emotion as the only emotion present	Range of percentages judging that both the relevant and the second most frequent emotion were both present
Happiness	10	10	91-97	0-1	0-6
Sadness	10	7	72-90	1-8	1-12
Surprise	10	8	64-84	1-4	1-18
Anger	10	8	50-71	3-13	12-31
Disgust	4	4	56-67	6-10	12-15
Fear	7	7	44-54	0-15	11-38

Three scorers were trained for about six hours each. Each FAST item was discussed, the relevant wrinkles or positions of the face in each item were indicated, important discriminations between particular pairs of items were emphasized. Practice photographs were then scored; discrepancies between the scores and those made by two of the authors (Paul Ekman and Wallace V. Friesen) were discussed. Scoring a photograph was found to be a fairly rapid procedure. A scorer usually scored a face on one facial area in less than thirty seconds.

The three scorers worked independently, with the set of FAST photographic items for a particular facial area on a board in front of them, so that all could be seen at once. Their task was to find the FAST item(s) which best matched the appearance of the face to be scored. The scorer had five choices:

(1) A single FAST item that best matched the facial area to be scored.

(2) Two FAST items, one indicated as a first choice, and another as a second choice, if there were two items that closely approximated the facial area to be scored.

(3) Two FAST items, indicated as tied scores for first choice, if there was asymmetry within the facial area to be scored (e.g., the left brow resembled one FAST item, and the right brow resembled another).

(4) A NEUTRAL score, if the facial area to be scored appeared to be in a normal or rest position or the wrinkles shown were inferred to be a permanent part of the physiognomy.

(5) A NO-SCORE, if there was some movement in the facial area to be scored but no FAST item that approximated its appearance.

As a last step, after all three facial areas were scored by the three scorers, the entire face was shown and scored on the head orientation code.

Facial area score

The three scorers did not always agree. (Reliability is reported later in "RESULTS".) An *a priori* procedure was devised for combining the results of the three independent scorers into a single set of scores for each facial area for each of the 51 faces. Each facial area was assigned nine points, which were to be distributed across the six emotion categories, the neutral category, and the no-score category. Each of the three scorers contributed three out of the nine points to a facial area. If a scorer indicated only one choice for a facial area, that choice received his three points regardless of whether the choice was a FAST item, neutral,

or no-score. If a scorer indicated both a first and second choice, the first choice was given two points and the second choice one point. And, if the scorer indicated a tied choice, each item received a point and a half. In assigning points, the FAST items chosen by each scorer were converted into the emotion categories the items represented. If a scorer made two choices but both were items for one emotion category, that category received all three of his points; (Table I shows more than one FAST item for certain emotion categories within certain facial areas).

Whole face emotion prediction

The points for each emotion category, and for the neutral and no-score categories were summed across the three facial areas for each of the 51 selected faces, to yield a total of 27 points for each face; (nine points from each of three facial areas). A simple set of *a priori* arbitrary rules was employed for determining from the distribution of points whether there would be no prediction, a single emotion prediction, or a blend prediction.

No prediction was made either if there were fewer than six points accumulated by any one emotion category, or if the no-score category had the highest number of points and the most frequent emotion category had accumulated fewer than nine points. A single emotion category was predicted either if a single category had two points more than the next most frequent category and the number of points for the predicted category was more than six, or if the neutral category had the most points, but there was an emotion category that had at least six points. If two categories had at least six points but were tied or within less than two points of each other, the head orientation code was consulted in an attempt to break the tie. If the head orientation score supported only one of the two tied emotion categories, then that was the predicted category. If the head orientation score supported both of the tied categories, or neither, then the blend of the two emotions was predicted.

RESULTS

Reliability

Table IV shows two measures of agreement among the scorers. The first measure is quite strict, requiring that the scorers agree exactly on their

choices. If, for example, two scorers had given the same first choice for a picture, but one of them had also given a second choice, this was counted as a disagreement. The table shows the percentage of the 51 faces on which two of the three and all three scorers agreed exactly. The

TABLE IV

Reliability on FAST Scoring
(Percent of the 51 Stimuli Scored Where Scorers Agreed)

	Exact agreement on both first and second Choice		First Choice of one Scorer is either first or second Choice of other Scorers	
	2 of 3 scorers	3 of 3 scorers	2 of 3 scorers	3 of 3 scorers
Brow-Forehead	80 %	31 %	98 %	69 %
Eyes	82 %	41 %	98 %	78 %
Lower face	88 %	59 %	98 %	78 %

second measure was a more generous index of agreement. Disagreement was noted only if one scorer's first choice was not either the first or second choice of another scorer. Again, agreement is shown for two and for all three of the scorers.

The extent of agreement shown in Table IV suggests there was sufficient reliability among scorers; the only low agreement figures are those measuring exact agreement among all three scorers. Examination of the pattern of agreement did not suggest that any one scorer was generally deviant. Instead, it appeared that, on occasion, any one scorer would fail to grasp all the details in the facial area to be scored, and thus would disagree with the other two scorers. The procedure for deriving affect scores for each facial area allows for such occasional deviance on scoring, since relatively few points are contributed to a prediction by any one scorer if his choices are not replicated by another scorer.

Validity

Table V shows that FAST correctly predicted the emotions judged by the observers for the majority of the faces. With 45 out of 51 faces correctly

TABLE V

FAST Prediction about Perceived Emotion in Relation to Emotion Judged by the Majority of Observers

Emotion judged by majority of the observers	FAST Prediction								
	Happiness	Surprise	Anger	Sadness	Disgust	Fear	Happiness Fear	Fear Surprise	Fear Surprise Sad
Happiness	10								
Surprise		10							
Anger			10						
Sadness				9			1		
Disgust				1	3				
Fear		1		1		3		1	1

predicted, there is no need for a statistical test to establish significance. FEAR was the only emotion category for which FAST scores did not succeed with the majority of the faces. Those errors were not totally irrelevant, since two of the four incorrectly predicted FEAR faces were predicted to be blends involving fear.

It may well be that FAST's relatively poor success with FEAR as compared to other emotion category faces was related to the fact that the FEAR faces were not as adequate in showing the single emotion as were the faces for the other emotion categories. Table II showed that when all three criteria utilized in selecting single emotion faces were considered, the FEAR pictures elicited less agreement among observers about that single emotion than did the other emotion category faces.

We believe that FAST would succeed as well in predicting FEAR as other emotion categories, if similarly high agreement, single emotion FEAR faces were scored. It might be argued, however, that perhaps such faces cannot occur. Since no such FEAR faces were found among all the faces initially considered, and there were high agreement faces for each of the other emotion categories, perhaps FEAR faces always elicit lower agreement or judgments of blended emotions. Yet, when we posed faces by having untrained persons imitate FAST FEAR items, those faces were judged with as much agreement about a single emotion as is shown in Table II for the emotion categories of SADNESS, SURPRISE and ANGER. However, until it can be shown that FAST can correctly predict the majority of some new set of faces which observers agree represent only FEAR, some question will have to remain about whether the FAST fear items are completely adequate for this task. (Our new FEAR faces could not be used for this purpose, since these faces were obtained by having persons imitate FAST photographic items).

Comparison of facial areas

Table VI shows the percentage of the 51 faces which were correctly predicted when FAST scores were considered for each of the three facial areas separately, as compared with the combined scores which were shown in Table V and which are shown again in Table VI. No one facial area provided as many correct predictions as did the combined facial area scores. Certain facial areas did, however, provide better predictions than other areas for particular emotions; and, for five emotion categories, a particular facial area did yield a prediction as good as or better than the combined facial area score.

TABLE VI

Comparison of Percent Correct Predictions From Each Separate Facial Area (FAST Scores)

Emotion category	Facial Area			
	Brows-Forehead	Eyes	Lower face	All three facial areas combined (data reported in Table V)
Happiness	70	90	100	100
Sadness	70	90	0	90
Surprise	70	90	90	100
Anger	80	50	100	100
Disgust	25	0	75	75
Fear	29	71	29	43
Correct predictions across all emotion categories	49	73	67	88

These findings contradict most of the results reported by earlier investigators who had attempted to determine whether one facial area was better than another by comparing observers' judgments made from different facial areas (Coleman, 1949; Dunlap, 1927; Hanawalt, 1944; Nummenmaa, 1964; Ruckmick, 1921). But there was little consistency among those investigators, probably because most chose to utilize the facial behavior of but one stimulus person. The data in Table VI were based on 28 different stimulus persons. It may be that particular people tend to employ one facial area more than another when attempting to pose a particular emotion. The present data do not allow study of that possibility, since we had selected very few faces from each stimulus person in order to have many different stimulus persons included. Table VI does suggest that certain facial areas may provide more information for one emotion than for another, and that there might be benefits in weighting scores from particular facial areas for predicting particular emotions. But replication of the differential importance of facial areas for particular emotions with another sample of stimulus persons should be obtained before reaching any such conclusions.

DISCUSSION

These results are very encouraging. It has been possible to specify and measure the particular facial behavior relevant to the recognition of six emotions. The extent of success achieved with this measurement procedure is remarkable, especially since it was based on *a priori* theory and hunches and employed simple, logically based rules for combining scores to yield predictions rather than weighted scores based on empirical results. FAST opens for investigation a large number of research questions previously neglected because of the lack of an adequate method for measuring facial behavior. FAST should serve to legitimate research on the face and emotion, since the stimulus basis for observers' impressions about emotions can now be specified. FAST should encourage studies where the facial behavior itself is the dependent variable, rather than observers' judgments of the face, where measurement of the face can be related to changes in environmental conditions, self-reported emotion, instructional set, etc.

It should be recognized, however, that the present report represents but a first step in establishing the validity of FAST, albeit a promising one. The necessary next steps for establishing validity, a few of the potential problems in this measurement procedure, and experiments we have in progress will now be discussed. The first validity test reported here employed still photographs of posed facial behavior judged as single emotion stimuli. We will first discuss the question of FAST's applicability to spontaneous behavior, to blends, and to motion rather than still records of the face. We will then consider FAST's use in expression rather than recognition studies, the problems of physiognomic variables, and the comprehensiveness of the scoring system.

Some investigators (e.g., Hunt, 1941) have viewed posed facial behavior as a unique set of culturally bound, learned conventions that have little relevance to spontaneous facial behavior. Since FAST succeeded in predicting the recognition of posed faces, that view might suggest that FAST would have little applicability to the study of spontaneous facial behavior. The fact that posed behavior is judged as showing the same emotions across eleven literate cultures and one preliterate culture (Ekman, 1968; Ekman and Friesen, 1970; Ekman, Sorenson, and Friesen, 1969; Izard, 1968, 1970a) argues against the view that posed behavior has no relationship to spontaneous facial behavior, since it is highly improbable that so many different cultures would have developed the same set of learned arbitrary conventions for each emotion. These

cross-cultural studies have provided strong evidence of a pan-cultural element in the facial behavior associated with emotion, namely that particular facial behaviors are associated with particular emotions. Significantly, association of the same emotions with the same faces across cultures was established with posed faces. Further, most of the facial stimuli used by Ekman *et al.* in their cross-cultural investigations were included in the FAST study reported here, so that the present results could be extended to show that FAST predicts how most of these faces will be judged by observers from a number of different cultures. In light of this evidence, it seems reasonable to assume that posed behavior differs little from spontaneous facial behavior in form. Ekman, Friesen, and Ellsworth (1971) have suggested that posed differs from spontaneous facial behavior in duration, in the lack of attempt to control or otherwise moderate the behavior, and in the frequency of single emotion faces as compared to blend faces. Direct evidence, however, is required to show that FAST can accurately predict how spontaneous faces are judged by observers. We have such a study in progress, and expect that it will suggest the need to add some items to FAST, since we may well have overlooked certain movements for some of the emotions.

Another validity test in progress is investigating whether FAST can discriminate faces judged as blends from those judged as single emotions, and the particular emotions that compose a blend. Blends can be shown on the face in three different ways: (1) division by facial areas, in which one emotion is shown in one facial area, another emotion shown in another facial area; (e.g., raised brows as in surprise, and pressed lips as in anger); (2) division within a facial area, in which one emotion is shown in one part of that facial area, another in another part of that area; (e.g., one brow raised as in surprise, the other lowered as in anger); (3) the display of a face that does not incorporate the elements of either blended emotion but instead is the product of the simultaneous muscle action relevant to each single emotion. This last type of blend face should not be accurately predicted by FAST, and might suggest the need for additional items. The discrimination of blends from single emotion faces will probably require the development of a more complex set of prediction rules utilizing some weighting of scores for each of the three facial areas rather than the simple summing of points and predicting the majority score that was used in the study reported here.

Another study in progress has been applying FAST to videotape records of facial behavior. FAST seems to be as applicable to such motion records as to still photographs, although the scoring of videotapes is

slower. Most of the excess time is required for determining the exact duration of each change in facial appearance, however, rather than determining which FAST item is applicable. The resolution required for FAST scoring of videotapes necessitates filling the videopicture with just the face. In one respect, FAST is easier with motion records than with still photographs. The scorer has no problem deciding whether a particular wrinkle is part of the permanent physiognomy of the person when utilizing motion records, since he has the base-line facial behavior and can discount anything that never changes in appearance.

FAST was designed to be applicable not only in recognition of emotion studies, but also in expression studies, where the aim is to predict the person's self-reported feelings or the conditions under which the facial behavior occurred. One such study is in progress utilizing the videotape records mentioned above. In a comparison of Japanese and U.S. subjects, FAST is being applied to spontaneous facial behavior emitted while the subjects watched stressful and neutral films (Ekman, Lazarus, Friesen, Opton, Averill, and Malmstrom, 1970). FAST seems sufficiently comprehensive to encompass most of the facial behavior shown in these videotapes, although certain facial behaviors, purposefully omitted from FAST because they were not considered emotion-specific, do occur in these records; (e.g., tongue showing, lip biting, cheek biting, etc.). Agreement among scorers is comparable to that reported earlier for the scoring of still photographs. Validity studies in progress will compare FAST scores with self-reported emotions, with psychophysiological measurements, and between the two experimental conditions, watching a neutral or a stress film.

Physiognomic determinants of facial behavior may have two undesired influences on FAST scores. People may differ in the permanent wrinkle patterns that with age become part of their facial appearance, more or less noticeable perhaps depending upon the action of the facial muscles. Some of those permanent wrinkles may resemble FAST items and not be discriminated from wrinkles produced by muscle movement. As mentioned earlier, this problem does not occur if motion records are utilized, so that the scorer can see the base line, unchanging wrinkles. With still photographs, the scorer may not be able to make that differentiation. This would not be a problem in a recognition study, since even permanent wrinkles are part of the stimulus configuration seen by the observer; if the permanent wrinkles are relevant to emotion, they should be scored by FAST. In expression studies, however, where the aim is to distinguish a transient event, a mood, an environmental change, etc., if FAST scores

the permanent wrinkles it is adding information irrelevant to that transient event. This may not be a serious problem, however, since in such expression studies the investigator would rarely decide to utilize a single still photograph as his record, but instead would utilize either multiple sequential stills or a motion record, and the scorer could see the base-line pattern.

A potentially more serious problem may result from individual differences in facial repertoire. Pilot studies requiring individuals to imitate FAST items suggest that there are reliable individual differences in ability to make certain facial movements. The distribution of such differences is not yet known, nor is it known whether they are due to bone structure, facial musculature, fatty deposits, or psychological factors. But if there are such individual lacunae in the repertoire of facial behavior included in FAST, they would weaken FAST predictions for particular people. For example, if a person is unable to produce vertical wrinkles between the brows and a lowering and drawing together of the inner corners of the brow, he cannot be eligible to receive a FAST anger score for the brow-forehead area. We have begun to specify rules that predict the six emotions even when certain elements within particular facial areas are not shown, on the basis of the presence of certain other items. In the example just described, if the lower eye lid movement and the mouth positions listed in FAST for ANGER were shown, but the brow resembled a DISGUST or SADNESS brow (which we have found can occur when some people attempt to imitate FAST's ANGER brow), or a neutral or no-score brow, the prediction would still be ANGER. Utilizing some of these decision rules based on such physiognomic considerations increased FAST's success in prediction. There were 49 out of 51 faces correctly predicted, rather than 45 out of 51, as reported earlier when the sum of scores across facial areas had been used. This suggests the value of this approach, but most of the work on determining such individual differences in ability to show the movements depicted in FAST and the development of relevant prediction rules remains to be done.

It is difficult to evaluate whether FAST is too elaborate a system, or too limited in its coverage. Considering only six emotion categories, FAST allows for more than 12,000 combinations of items for the three facial areas. Obviously, some of these must not be anatomically possible. In studying more than 2,000 of the combinations, however, we have not yet found one of the impossible ones. There may well be more items necessary for these six emotions, and this should be discovered in the spontaneous recognition and spontaneous expression studies in progress.

FAST will be broadened to cover emotions of INTEREST, CONTEMPT, and SHAME. Other emotions might be considered, depending upon the theoretical orientation of the investigator. We believe that other emotions beyond these mentioned either do not have a distinctive facial appearance or can be conceived as blends of those we have considered, but there is no conclusive research as yet to support that belief.

While FAST presently allows for thousands of facial appearances, it should be remembered that it was designed to describe and measure facial behaviors associated with emotion. Many facial behaviors that we considered as either not emotion-specific or not relevant to emotion (e.g., tongue showing, lip biting, etc.) are not included in FAST, and presumably these nonemotional facial behaviors far outnumber the facial appearances covered by FAST.

REFERENCES

Allport, F.H.,
1924 *Social Psychology* (Boston: Houghton Mifflin).

Arnold, M.B.,
1960 *Emotion and Personality* (New York: Columbia University Press).

Birdwhistell, R.L.,
1952 *Introduction to Kinesics* (Louisville, Ky.: University of Louisville Press).

Coleman, J.C.,
1949 "Facial Expressions of Emotion", *Psychological Monograph* 63:296.

Darwin, Charles,
1872 *The Expression of the Emotions in Man and Animals* (London: John Murray).

Davis, R.C.,
1934 "The Specificity of Facial Expressions", *Journal of General Psychology* 10, 42-58.

Duchenne, B.,
1862 *Mécanisme de la physionomie humaine ou analyse électro-physiologique de l'expression des passions* (Paris: Baillière).

Dunlap, K.,
1927 "The Role of Eye-Muscles and Mouth-Muscles in the Expression of the Emotions", *Genetic Psychology Monograph* 2, 199-233.

Ekman, P.,
1968 "The Recognition and Display of Facial Behavior in Literate and Non-Literate Cultures", *Proceedings, American Psychological Association Convention* 3, 727.

Ekman, P., and W.V. Friesen,
1968 "Nonverbal Behavior in Psychotherapy Research", *Research in Psychotherapy* 3, ed. by J. Shlien (Washington, D.C.: American Psychological Association), 179-216.
1969a "Nonverbal Leakage and Clues to Deception", *Psychiatry* 32:1, 88-105.
1969b "The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding", *Semiotica* 1, 49-98.

1970 "Constants across Culture in the Face and Emotion", *Journal of Personality and Social Psychology* (in press).

Ekman, P., W.V. Friesen, and P. Ellsworth,
1971 *The Face and Emotion: Guidelines for Research and an Integration of Findings* (New York: Pergamon).

Ekman, P., R.S. Lazarus, W.V. Friesen, E.T. Opton, J.R. Averill, and E.J. Malmstrom,
1970 "Facial Behavior and Stress in Two Cultures", unpublished manuscript (Langley Porter Neuropsychiatric Institute).

Ekman, P., E.R. Sorensen, and W.V. Friesen,
1969 "Pan-Cultural Elements in Facial Displays of Emotions", *Science* 164, 86-88.

Engen, T., N. Levy, and H. Schlosberg,
1957 "A New Series of Facial Expressions", *American Psychologist* 12, 264-66.

Frijda, N.H.,
1969 "Recognition of Emotion", *Advances in Experimental Social Psychology* 4, ed. by L. Berkowitz (New York: Academic Press).

Frijda, N.H., and E. Philipszoon,
1963 "Dimensions of Recognition of Emotion", *Journal of Abnormal Social Psychology* 66:1, 45-51.

Frois-Wittmann, J.,
1930 "The Judgment of Facial Expression", *Journal of Experimental Psychology* 13, 113-51.

Fulcher, J.S.,
1942 " 'Voluntary' Facial Expression in Blind and Seeing Children", *Archives of Psychology* 38:272.

Hanawalt, N.G.,
1944 "The Role of the Upper and the Lower Parts of the Face as the Basis for Judging Facial Expressions II: In Posed Expressions and 'Candid Camera' Pictures", *Journal of General Psychology* 31, 23-36.

Honkavaara, S.,
1961 "The Psychology of Expression", *The British Journal of Psychology*, Monograph Supplements 23, ed. by R.H. Thouless (New York: Cambridge University Press).

Hunt, W.A.,
1941 "Recent Developments in the Field of Emotion", *Psychological Bulletin* 38:5, 249-76.

Izard, C.,
1968 "Cross-Cultural Research Findings on Development in Recognition of Facial Behavior", *Proceedings, American Psychological Association Convention* 3, 727.
1970a "The Emotions and Emotion Constructs in Personality and Culture Research", *Handbook of Modern Personality Theory*, ed. by R.B. Cattell (Chicago: Aldine).
1970b "The Emotions as a Culture-Common Framework of Motivational Experiences and Communicative Cues", *The Subjective Culture*, ed. by G. Vassiliou and V. Vassiliou (in press).

Landis, C.,
1924 "Studies of Emotional Reactions II: General Behavior and Facial Expression", *Journal of Comparative Psychology* 4, 447-509.

Landis, C., and W.A. Hunt,
1939 *The Startle Pattern* (New York: Farrar).

Leventhal, H., and E. Sharp,
1965 "Facial Expressions as Indicators of Distress", *Affect, Cognition and Personality: Empirical Studies*, ed. by S.S. Tomkins and C.E. Izard (New York: Springer), 296-318.

Murphy, G., L.B. Murphy, and T.M. Newcomb,
1937 *Experimental Social Psychology*, rev. ed. (New York and London: Harper Brothers).

Nummenmaá, T.,
1964 *The Language of the Face (= Jyväskylä Studies in Education, Psychology, and Social Research)* (Jyväskylä, Finland: Jyväskylän Yliopistoyhdystys).

Plutchik, R.,
1962 *The Emotions: Facts, Theories, and a New Model* (New York: Random House).

Ruckmick, C.A.,
1921 "A Preliminary Study of the Emotions", *Psychology Monograph* 30:134-39, 30-35.

Thompson, J.,
1941 "Development of Facial Expression of Emotion in Blind and Seeing Children", *Archives of Psychology* 37:264.

Tomkins, S.S.,
1962 *Affect, Imagery, Consciousness*, Vol. 1: *The Positive Affects* (New York: Springer).
1963 *Affect, Imagery, Consciousness*, Vol. 2: *The Negative Affects* (New York: Springer).

Tomkins, S.S., and R. McCarter,
1964 "What and Where Are the Primary Affects?: Some Evidence for a Theory", *Perceptual and Motor Skills* 18, 119-58.

Trujillo, N.P., and T.A. Warthin,
1968 "The Frowning Sign Multiple Forehead Furrows in Peptic Ulcer", *Journal of the American Medical Association* 205:6, 218.

Paul Ekman (born 1934) is Professor of Psychology in Residence at the Department of Psychiatry, School of Medicine, University of California, San Francisco; he is presently Principal Investigator on three Research Grants on Studies of Nonverbal Behavior. His major publications are: "Head and Body Cues in the Judgment of Emotion: A Reformulation" (1967); "Nonverbal Behavior in Psychotherapy Research" (1968); "VID-R and SCAN: Tools and Methods in the Analyses of Facial Expression and Body Movement" (1969); "Nonverbal Leakage and Clues to Deception" (1969).

Wallace V. Friesen (born 1933) is Co-Principal Investigator on two Research Grants on Studies of Nonverbal Behavior, and co-author with Paul Ekman of the above works.

Silvan Tomkins (born 1911) is a Professor in the Department of Psychology at Livingston College, Rutgers — The State University. The principal focus of his research is personality theory. Among his chief publications are: *Thematic Apperception Test* (1947); *Affect, Imagery, Consciousness*, 1 (1962) and 2 (1963); *Affect, Cognition, and Personality*, ed. with C. Izard (1965).