

- Covington, M., & Omelich, C. (1979b). It's best to be able and virtuous too: Student and teacher evaluative responses to successful effort. *Journal of Educational Psychology*, 71, 688-700.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192-238). Lincoln: University of Nebraska Press.
- Kelley, H. H. (1971). *Attribution in social interaction*. Morristown, NJ: General Learning Press.
- Kelley, H. H. (1972). *Causal schemata and the attribution process*. Morristown, NJ: General Learning Press.
- Kukla, A. (1972). Foundations of an attributional theory of performance. *Psychological Review*, 79, 454-470.
- Lanzetta, J., & Hannah, T. (1969). Reinforcing behavior of "naive" trainers. *Journal of Personality and Social Psychology*, 11, 245-252.
- Meyer, W.-U. (1973). *Leistungsmotiv und Ursachenerklärung von Erfolg und Misserfolg* [Achievement motive and causal explanation of success and failure]. Stuttgart, Federal Republic of Germany: Klett.
- Meyer, W.-U. (1978). Der Einfluss von Sanktionen auf Begabungspereptionen [The influence of feedback on perceived ability]. In D. Görlitz, W.-U. Meyer, & B. Weiner (Eds.), *Bielefelder Symposium über Attribution* [Bielefeld Symposium on Attribution] (pp. 71-87). Stuttgart, Federal Republic of Germany: Klett-Cotta.
- Meyer, W.-U. (1982). Indirect communications about perceived ability estimates. *Journal of Educational Psychology*, 74, 888-897.
- Meyer, W.-U. (1984). *Das Konzept von der eigenen Begabung* [Self-concept of ability]. Bern, Switzerland: Huber.
- Meyer, W.-U., Bachmann, M., Biermann, U., Hempelmann, M., Plöger, F.-O., & Spiller, H. (1979). The informational value of evaluative behavior: Influences of praise and blame on perceptions of ability. *Journal of Educational Psychology*, 71, 259-268.
- Meyer, W.-U., Engler, U., & Mittag, W. (1982). Auswirkungen von Tadel auf die Beurteilung des eigenen Leistungsstandes und auf das Interesse an Aufgaben [Effects of blame on the assessment of one's own performance and task interest]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 14, 263-276.
- Meyer, W.-U., & Plöger, F.-O. (1979). Scheinbar paradoxe Wirkungen von Lob und Tadel auf die wahrgenommene eigene Begabung [Seemingly paradoxical effects of praise and blame on perceived own ability]. In S.-H. Filipp (Ed.), *Selbstkonzept-Forschung. Probleme, Befunde, Perspektiven* [Research on self-concept: Problems, results, perspectives] (pp. 221-235). Stuttgart, Federal Republic of Germany: Klett-Cotta.
- Plöger, F.-O. (1980). *Sanktionen und Begabungswahrnehmung: Alltagspsychologische und entwicklungspsychologische Aspekte* [Feedback and perceived ability: General and developmental aspects]. Unpublished master's thesis, University of Bielefeld, Federal Republic of Germany.
- Rest, S., Nierenberg, R., Weiner, B., & Heckhausen, H. (1973). Further evidence concerning the effects of perceptions of effort and ability on achievement evaluation. *Journal of Personality and Social Psychology*, 28, 187-191.
- Weiner, B., & Kukla, A. (1970). An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology*, 15, 1-20.

FUZZY CONCEPTS AND THE PERCEPTION OF EMOTION IN FACIAL EXPRESSIONS

JAMES A. RUSSELL

University of British Columbia

MERRY BULLOCK

Max-Planck-Institute for Psychological Research

Research on how well one person can recognize the emotion expressed in another person's face has resulted in controversy: accuracy versus inaccuracy; discrete categories versus dimensions versus structural models of emotion. These seemingly disparate conclusions can be reconciled when natural language concepts of emotion ("happiness," "anger," "fear," "sadness," etc.) are thought of as overlapping and fuzzy, rather than as mutually exclusive and properly defined. The research literature on the perception of emotion in facial expressions is reviewed from this vantage point, and four studies testing predictions from this thesis are reported. When rating the degree to which either posed or spontaneous facial expressions exemplify emotion categories, subjects produced reliably graded responses and indicated that individual (even prototypical) expressions belong to more than one category. The graded "prototypicality" ratings (1) predicted the probability with which the expression was said to be a member of the emotion category or was selected as a referent for the category label, and (2) generated a replica of a circumplex structural model of emotion when subjected to multidimensional scaling. Methodological and conceptual issues about the meaning attributed to facial expressions and about the perception of emotion in general are also discussed.

In this paper, we discuss the ordinary language concepts that people use to distinguish types of emotion: concepts such as "anger," "fear," "happiness," and so on. Our general aim is to suggest that

We thank Katherine Bodie and Lorraine Nelos for their help on this study, and Larry Ward for his comments on an earlier draft of this paper. This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada. Requests for reprints should be sent to James A. Russell, Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Y7.

viewing these concepts as fuzzy, rather than as properly defined, can help clarify some of the traditional issues in the psychology of emotion. We illustrate this claim here by considering controversies surrounding the perception of emotion in facial expressions.

Writers in various disciplines now argue that many natural language concepts are not properly defined (James, 1902/1929; Labov, 1973; Lakoff, 1973; Langer, 1942/1980; Mervis & Rosch, 1981; Wittgenstein, 1953; Zadeh, 1965). Rosch's (1973, 1975, 1977) research on natural language categories of objects revealed features of such concepts as "vehicle" and "furniture" not anticipated by the classical view in which concepts were thought to be understood by means of necessary and sufficient features. Her prototype theory of concepts has inspired a singularly fruitful re-examination of such important psychological concepts as the self, personality traits, psychiatric diagnostic categories, social stereotypes, intelligence, and social roles (Brown, 1980; Buss & Craik, 1983; Cantor & Mischel, 1977, 1979; Cantor, Smith, French, & Mezzich, 1980; Dahlgren, 1985; Hampson, 1982; Horowitz, Wright, Lowenstein, & Parad, 1981; Neisser, 1979; Rogers, Kuiper, & Kirker, 1977; Wiggins, 1980).

Rosch's is not the only nonclassical theory, and it is not yet clear how best to characterize everyday concepts (Jones, 1982; Mervis & Rosch, 1981; Osherson & Smith, 1981; Smith & Medin, 1981), or even whether one characterization will apply to all (Armstrong, Gleitman, & Gleitman, 1983). Our argument, however, rests on just a few key interrelated properties that do appear to contrast many everyday concepts with properly defined ones. In a properly defined category, something either is or is not a member; all members are equal in membership; and certain sets of properly defined categories are mutually exclusive. To take geometric figures as an example, any figure either is a square or is not; all squares are equally square; and the categories "square," "triangle," and "circle" are mutually exclusive. When we say here that emotion categories are fuzzy, we mean that they lack these three properties—that, in the domain of emotion categories, borders between categories are fuzzy rather than clear-cut; membership is a matter of degree; and different categories tend to overlap one another.

It is helpful to begin by picturing the domain of everyday emotion concepts as a hierarchical structure, which is a metaphorical way of saying that some categories are included within others. At the top, most inclusive level is the superordinate concept of "emotion." At the middle level, emotion is divided into various categories ("love," "fear," "sadness," etc.), at least some of which may be further subdivided ("love" into "romantic love," "filial love," etc.), forming a

subordinate level in the hierarchy. The superordinate concept of emotion appears to be fuzzy. For example, anger, fear, and love are prototypical examples of emotion; pride and embarrassment are somewhat poorer examples; and calmness, respect, and sleepiness may or may not be emotions (Bullock & Russell, 1984; Fehr & Russell, 1984; Fehr, Russell, & Ward, 1982; Russell, 1980). The present paper is concerned with the idea that middle-level categories of emotion are also fuzzy.

Elsewhere, we consider more direct semantic tests of this idea (Russell, 1984). Here, we focus on its *usefulness*: Does assuming that middle-level emotion concepts are fuzzy, in the sense just defined, provide a vantage point from which, for example, we can better understand how a person uses these concepts to interpret the emotion of another? We concentrate here on how a person interprets another's facial expression, and we argue that the simple theoretical tack of thinking of emotion concepts as fuzzy throws light on such controversies as the accuracy with which people read facial expressions, discrete categories versus dimensions of emotion, and the meaning of structural models of emotion. We first attempt to show that the literature on the perception of emotion in facial expressions can be profitably reinterpreted from this perspective. We next report four studies that provide preliminary empirical support for our interpretation of that literature. Finally, we extend our analysis to other issues concerning the perception of emotion.

PERCEPTION OF EMOTION IN FACIAL EXPRESSIONS

THREE TRADITIONAL POSITIONS

The modern era of research on how emotion is read from the face began with Charles Darwin's (1872/1965) argument that the human face provides information about emotion that is easily recognized by another human. Many scientists agreed with Darwin, but others have remained skeptical of this idea. Ekman, Friesen, and Ellsworth (1982) have claimed that "the question of whether the face can provide accurate information about emotion has been the central issue since the beginning of research on the face" (p. 56).

Any direct assessment of "accuracy" would require knowledge of the true state of affairs and is probably an intractable problem. Most researchers therefore assessed something that seemed related to accuracy: the *consistency* with which people *categorize* (label) the emotional meaning of a facial expression. "Consistency" has meant con-

sistency with the experimenter's label; with the modal response from other subjects; or, in the case of a posed facial expression, with the emotion intended by the actor who posed it. At least one result of this research is undeniable: "academic storm and strife over the years" (Schneider, Hastorf, & Ellsworth, 1979, p. 207). Three apparently conflicting positions have been taken.

The first position stems from studies carried out in the early decades of this century in which researchers repeatedly found *inconsistency* when subjects labeled facial expressions. When situational cues were eliminated, different subjects often gave different labels for the same facial expression (Hunt, 1941; Landis, 1924, 1929; Sherman, 1927a, 1927b), and subjects' labels appeared to be easily influenced by suggestion (Fernberger, 1928). Reviewing a half-century's accumulated research, Bruner and Tagiuri (1954) concluded that it was doubtful that emotion could be accurately read from the face alone. This first position appears today in introductory textbooks when the reader is asked to observe how a face alone is ambiguous for signaling a specific emotion and how the context is needed to disambiguate what the emotion is. Although the decades-old research that generated this idea has been criticized for its methodological shortcomings (Woodworth & Schlosberg, 1954), more recent and methodologically sophisticated research can be cited in its support (Spignesi & Shor, 1981; Wagner, MacDonald, & Manstead, 1986).

The second position is that people are not all that inconsistent, once a different criterion for consistency is adopted. Woodworth (1938) argued that some inconsistencies are more apparent than real because facial expressions of different emotions vary in how similar they are to one another. Examining "errors" more closely, Woodworth found that they follow a systematic pattern, from which he derived a simple structural model of emotion. Woodworth ordered emotions along a continuum such that adjacent emotions were the ones most likely to be "confused" with one another—that is, labeled in a similar fashion. Woodworth's student, Schlosberg (1952), noticed that the ends of Woodworth's continuum were also occasionally "confused" with each other and therefore tied the two ends together, forming his well-known circular structural model of emotions. In turn, the circular order suggested two underlying dimensions, one reasonable interpretation of which seems to be degree of pleasure and degree of arousal (e.g., Abelson & Sermat, 1962; Russell & Bullock, 1985, 1986). In short, Woodworth and Schlosberg believed that people are consistent—provided that consistency is defined as placement within a region of an emotional structure defined by dimensions, rather than as placement into categories described by the terms of our every-

day language. This work led to that current theorizing that emphasizes dimensions and structural models of emotion over everyday categories (e.g., Mehrabian, 1972; Russell, 1980).

The third position is that people do read the emotional messages of facial expressions consistently and do so in terms of basic, everyday emotion categories such as "anger," "happiness," "fear," and "surprise." This position derives from a series of studies (Ekman, Sorenson, & Friesen, 1969; Izard, 1971; Tomkins & McCarter, 1964) stimulated by Tomkins's (1962–1963) theory of emotion. These researchers showed that certain facial expressions are categorized with very high (although still not perfect) agreement, even by members of different cultures. Unlike the earlier researchers, the Tomkins group used carefully selected facial expressions and restricted their attention to a small number of "basic" emotions: Ekman *et al.* (1969) examined happiness, sadness, anger, fear, disgust, and surprise. To these, Izard (1971) added interest and shame.

A REINTERPRETATION

"Academic storm and strife" among these three positions may be unnecessary. Viewing emotion categories as fuzzy allows us to outline a coherent account that integrates the various findings on which the positions are based.

What the Tomkins group demonstrated, we believe, is that there exist *prototypical* expressions for certain emotion categories. Recall that these researchers had searched for those particular facial expressions that best represent an emotion category. Ekman and Friesen (1976) describe how they searched through hundreds of photographs to find just the right ones. The best representatives are the prototypical representatives. In our view, 100% agreement among subjects labeling a facial expression is one index of its prototypicality, its full degree of membership, in the selected category. With membership measured on a scale from 0 to 1, these are the ones.

What the early researchers who reported inconsistency demonstrated, we believe, is that some facial expressions are less prototypical members of a category (membership less than 1) and are members of more than one category. The expressions they studied were often realistic or spontaneous rather than dramatized prototypes. Rather than view subjects' judgments as "errors," we take them as an indication that more than one label is indeed applicable to the same facial expression. For example, if a particular expression has an intermediate and equal degree of membership (e.g., .5) in two

categories, then a choice between the two corresponding emotion labels is arbitrary. Forced to choose, half the subjects may select one label, half the other. Thus, the .5 degree of membership results in approximately 50% endorsement.¹

What Woodworth and Schlosberg demonstrated, we believe, is that the overlap of middle-level emotion categories is systematic. Although Woodworth and Schlosberg maintained the terminology of "accuracy," "correct" responses, and "errors" derived from the traditional view of concepts, their analysis of "errors" is more compatible with fuzzy boundaries and degrees of membership. Indeed, from our point of view, the fuzzy boundaries between categories are what allowed their structure to be discovered. More important, the fuzziness of emotion categories helps interpret the nature of the structure they discovered.

The Woodworth-Schlosberg structure (and modern variants of it—see Daly, Lancee, & Polivy, 1983; Russell, 1980; Zevon & Tellegen, 1982) represents a set of relationships among middle-level categories: an intercategory structure. It is a simultaneous representation of the ways in which the middle-level categories overlap each other in their extension in the domain of facial expressions. From the classical view of concepts, it was reasonable to assume that two words either label the same category and are therefore synonyms, or label distinct categories and are therefore mutually exclusive. From our view of concepts, in contrast, emotion words label categories that can show any degree of similarity (overlap) of referential meaning. Various measures of similarity are possible: For example, one could calculate the correlation between the degree-of-membership ratings for the two categories across a broad set of emotional events. Two labels for the same category (synonyms) would (with error-free data) be perfectly correlated. Other categories might be highly correlated, some moderately correlated, some uncorrelated, and some negatively correlated to varying degrees.

In summary, "academic storm and strife" resulted from common

1. This idea can be generalized to predict a positive correlation between degree of membership in a category and the proportion of subjects labeling it with the category name. Consider a subject who is shown a facial expression and asked to choose a single label from a list of labels. Presumably, the subject first assesses its degree of membership in each of the categories listed, and then chooses the label for the category achieving the highest membership value. If we allow random error in assessing degree of membership, then there will be variation in the label chosen whenever degree of error exceeds the differences between "true" membership values. The frequency of choice of each label for an expression would be roughly proportional to its degree of membership in that category.

but unstated assumptions about the nature of everyday language categories of emotion. To ask subjects to select one label from a set of six category labels may have presupposed that each facial expression was a member of one and only one category. There was one right answer; all else was error. When subjects failed to act accordingly, their behavior was difficult to interpret, and arguments arose over the best interpretation. The empirical findings from the resulting three camps become consistent and understandable when viewed in an alternative way. Membership in an emotion category is a matter of degree, and the border between categories is fuzzy. Facial expressions tend to be members of several categories, although to different degrees. Overlapping membership is systematic, revealing an intercategory structure and underlying dimensions.

OVERVIEW OF FOUR STUDIES

At this point, it will be helpful to examine evidence on the feasibility of the ideas outlined so far. Obviously, we could not study all the natural language categories for emotional states communicated via facial expression. Instead, we focused on those categories writers on emotion have most commonly referred to as "basic" (categories such as "fear," "anger," and "happiness"), but we also included categories less prototypical of emotion, (such as "calmness") and borderline categories (such as "sleepiness"). In our first study, subjects rated the degree of membership of various facial expressions in each of these categories. We sought to establish that each category exhibits graded and overlapping membership, and that together the categories exhibit an intercategory structure. We could then test our hypothesis on the relationship of the membership function of each category to the intercategory structure of the entire set (specifically, to the model of intercategory structure proposed by Russell, 1980).

Three subsequent studies replicated and extended the results of Study 1. In Study 2, we examined the criterion for fuzzy borders: subjects' inability to adjudicate membership in specific emotion categories. We also asked whether the probability that a facial expression would be endorsed as a member of a category could be predicted from its rated degree of membership in that category. In Study 3, we asked whether degree-of-membership ratings could also predict responses to a traditional task designed to test whether subjects can "recognize" members of emotion categories. In Study 4, we looked at judgments about spontaneous expressions of genuine feelings.

STUDY 1: DIRECT RATINGS OF DEGREE OF MEMBERSHIP

Subjects can readily rate the degree to which items are members of such categories as "furniture" and "vehicle"; what is more important, subjects tend to agree with one another in their ratings (McCloskey & Glucksberg, 1978; Rosch, 1973). In our first study, we sought similar evidence for middle-level categories relevant to emotion. Subjects were asked to rate 14 facial expressions for their degree of membership in each category.

In selecting facial expressions, we sought prototypical expressions of "basic" emotions. It would hardly do to demonstrate the fuzziness of emotion concepts with blurry photographs of ambiguous expressions or expressions of "blends" of several emotions. Fortunately, in *Pictures of Facial Affect*, Ekman and Friesen (1976) published sharp photographs of prototypical expressions of what they consider basic emotions. We selected examples from their set to provide a conservative test for our hypotheses.²

METHOD

Subjects

Subjects were 20 undergraduates at the University of British Columbia who were paid for their participation.

Stimuli

We chose 14 facial expressions that appeared to be the best representatives of certain emotion categories. Nine of the facial expressions were taken from Ekman and Friesen's (1976) *Pictures of Facial Affect*; these represent prototypical expressions of Ekman's (1972) list of six basic emotion categories. From our own collection were added five other photographs of actors asked to pose "excitement," "calmness," and "sleepiness." Expressions of these last three categories were needed to provide a more complete representation of the structural model of

2. This set also allowed us to maintain some distinction between prototypicality and intensity. In some domains, intensity plays no role in prototypicality: An apple is no more intense than a fig, but the apple is more prototypical of the concept "fruit" than is the fig. In other domains, intensity is one of the features that determines prototypicality: An intense red is a more prototypical red than is a dull red. Emotion is likely to be a domain of the second type: Intense anger is more prototypical than mild anger. In the four studies here, we sought to establish systematic and reliable differences in prototypicality, with intensity held approximately constant.

emotions to be tested. The expressions were labeled A through N. The categories they were intended to represent were as follows: A, excitement; B, happiness; C, happiness; D, calmness; E, calmness; F, sleepiness; G, sleepiness; H, sadness; I, sadness; J, disgust; K, anger; L, fear; M, surprise; and N, surprise. Both our and Ekman's emotion labels for the expressions are merely convenient devices and play no role in this study. The letters A through N serve as a reminder that a single emotion label for an expression is inconsistent with the ideas proposed here and to distinguish the facial expressions from emotion-denoting terms. (The order A through N corresponds to the order predicted by our intercategory structure; see Russell, 1980.)

Procedure

Subjects read the following instructions, which were adapted from Rosch (1973):

This study has to do with what we have in mind when we hear and use words. Consider the word "red." Close your eyes and imagine a true red. Now imagine an orangish red. Imagine a purplish red. Although you might still name the orange-red or the purple-red with the term "red," they are not as good examples of "red" (not as clear cases of what "red" refers to) as the clear true red. Orange and purple are even poorer examples of "red," perhaps not even red at all.

Notice that to judge how good an example something is has nothing to do with how much you like the thing. You might prefer a purple-red or orange-red to a true red, but still recognize which is a better example of "red."

The words we are interested in refer to feelings, moods, or emotions. We are interested in which facial expressions are better or poorer at conveying certain emotions.

The subjects were then given one term from the list in Table 1 and were asked to rate each of 14 facial expressions as examples of that concept. The ratings were made on a scale from 0 ("not an example at all") to 6 ("extremely good example"). This procedure was repeated for seven different emotion labels from the list for each subject. The order of presentation of emotion labels and facial expressions was random.

RESULTS AND DISCUSSION

Internal Structure

Mean prototypicality ratings are given in Table 1, with split-half estimates of the reliability of these ratings in the last column. Reading across the rows in Table 1 shows that, for each category, expressions

TABLE 1
Degree of Membership of Emotional Facial Expressions in Emotion Categories

CATEGORY	PHOTOGRAPH OF FACIAL EXPRESSION														RELIABILITY
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
Excited	4.7	3.6	1.6	1.2	0.9	0.3	0.2	0.4	1.8	1.2	2.4	3.0	2.7	3.6	.90
Happy	5.7	5.6	5.1	3.3	2.9	1.8	1.5	0.3	0.2	0.5	0.1	0.2	0.6	2.7	.92
Glad	5.3	5.8	4.5	4.0	3.3	1.6	0.9	0	0	0.3	0.1	0	0.7	2.8	.97
Calm	1.5	3.9	5.0	5.6	5.6	4.9	4.5	2.2	0.3	0.6	0.1	0	0.8	1.2	.96
Sleepy	0.1	0.3	0.1	1.5	3.4	1.9	4.8	2.4	1.3	0.3	0.1	0.3	0.3	0.1	.86
Sad	0.1	0.1	0.2	0.5	1.9	2.5	3.4	5.5	5.2	0.8	0.2	1.3	0.5	0.1	.96
Unhappy	0	0	0.4	0.6	1.4	2.5	3.4	5.8	5.6	3.8	4.2	3.5	1.8	0.7	.96
Miserable	0	0	0.2	0.6	1.0	2.6	3.1	5.6	5.9	3.7	3.7	3.1	1.2	0.6	.89
Disgusted	0.1	0.1	0.4	0.2	1.2	2.0	1.2	1.4	0.9	5.2	2.6	1.1	1.0	0.2	.73
Mad	0.2	0.2	0.1	0.4	0.7	1.4	1.0	1.6	1.9	3.2	5.4	2.5	1.2	0.3	.88
Angry	0	0	0.2	0.4	0.3	1.3	0.9	1.9	1.8	4.6	5.3	2.3	0.4	0.2	.95
Scared	0	0	0	0	0.2	0	0.3	1.4	2.6	1.0	1.1	5.5	4.8	1.8	.97
Afraid	0.3	0	0.3	0.2	0.1	0.4	0.3	0.9	3.1	0.9	1.5	5.3	3.6	1.9	.96
Surprised	5.5	1.7	0.9	0.4	0.2	0.2	0.1	0.6	1.1	1.3	1.5	3.6	5.8	4.9	.96

Note. Degree-of-membership ratings are means ($n=10$) on a scale from 0 ("not an example at all") to 6 ("extremely good example").

varied in the degree to which they exemplified that category—each category showed what is called an "internal structure." The last row, for example, gives results for the category "surprised." The peak, or best example, of "surprised" was Ekman's prototypical surprise expression, picture M. Another Ekman surprise photograph, N, was also an excellent example, although no better than photograph A, which showed someone attempting to appear excited. Ekman's prototypical expression of fear, expression L, was a moderately good example of "surprised." Ekman's prototypical expression of happiness, expression B, and his prototypical expression of anger, expression K, were borderline members of "surprised." Some other expressions were poor examples of "surprised."

Examination of frequency distributions of the raw scores showed that gradedness was not produced by averaging across a bimodal distribution: Individual subjects often produced intermediate ratings.

Intercategory Structure

Table 1 also shows that categories overlapped. Overlap is illustrated in Figure 1, which shows the mean degree of membership for three categories: "sad," "disgusted," and "angry." Although the peaks in the figure differ (and correspond to Ekman's prototypical expressions), each category label was, to different degrees, applied to some of the same expressions. The categories "angry" and "disgusted" overlapped to a considerable extent—a finding consistent with Ekman's (1972) cross-cultural finding that these two are often "confused" with each other.

Notice that in Figure 1 (and in Table 1), the expressions appear to be systematically ordered. The curves are not haphazard, but rise gradually to a peak and then descend. This sort of gradual curve originally gave rise to the name "fuzzy set." This systematic order within an emotion category would not be discernible if the 14 facial expressions were ordered along the abscissa of Figure 1 randomly. The gradual ascent and descent of the boundaries of such concepts as "red" or "tall" are visible because the stimuli can be arrayed along some physical dimension—colors plotted by wave length or objects by height. The internal structure of emotion concepts is discernible only because the photographs were ordered *a priori* in a certain way.

The ordering along the abscissa of Figure 1 was derived from a structural model of emotions (Russell, 1980), which places emotion categories in a circular order. We "clipped" the circle at a convenient spot and spread out the 14 expressions in a single line, ordered A through N. To test this *a priori* ordering against the present data, we

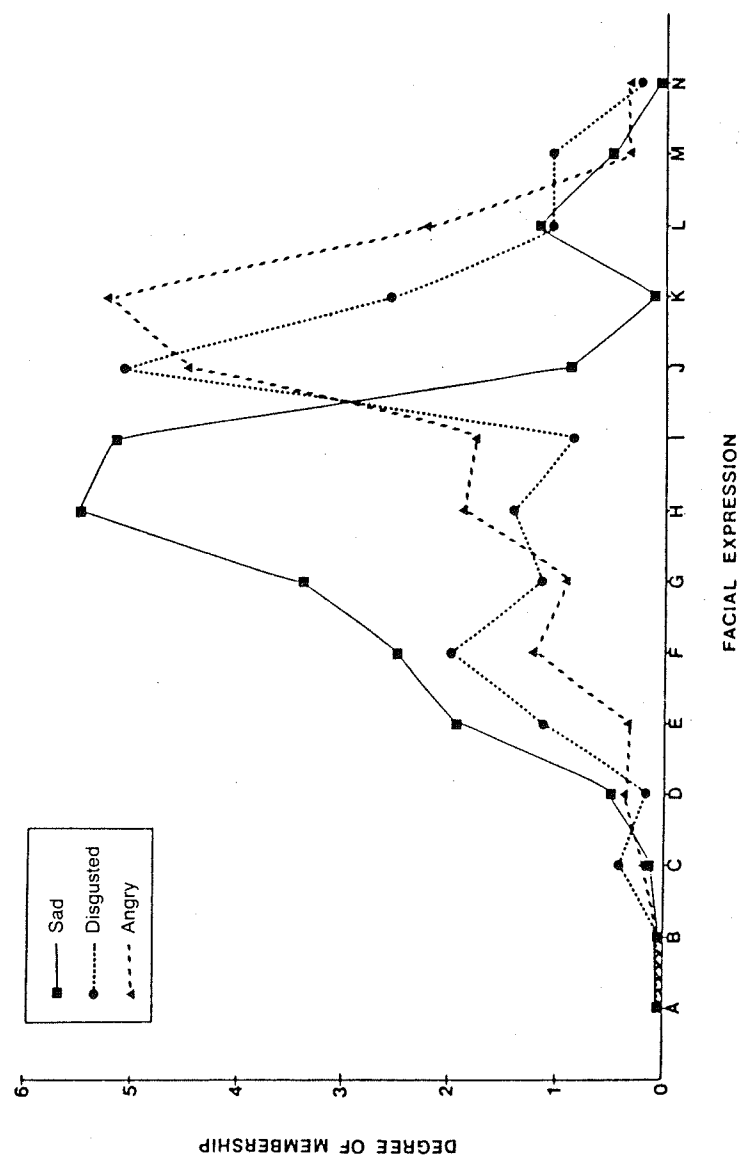
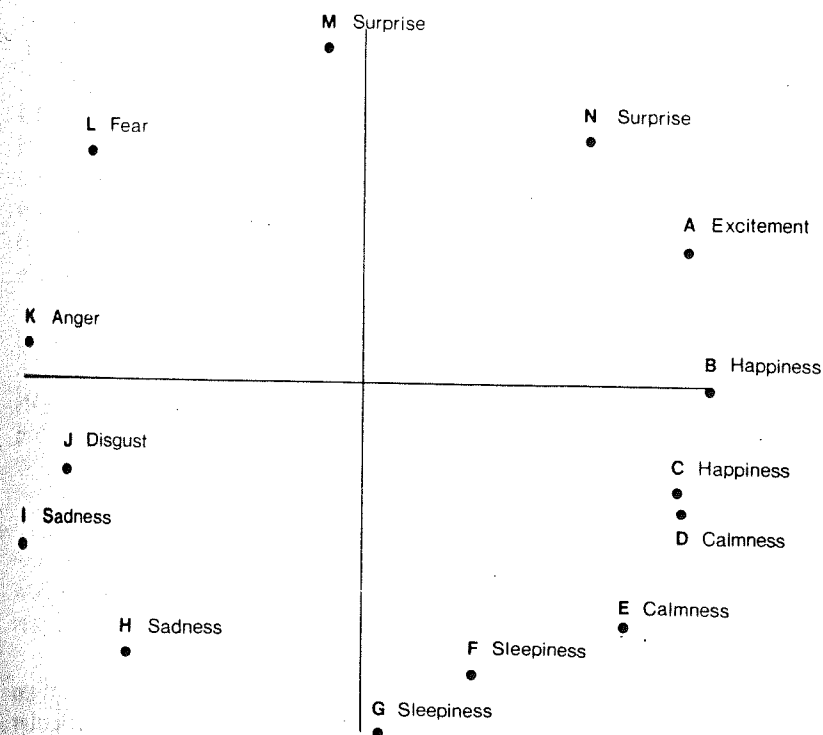


FIGURE 1
Internal structure of the categories of "sad," "disgusted," and "angry." Letters refer to different facial expressions.

derived a structural representation of the emotional facial expressions from the prototypicality ratings. Our measure of similarity between a pair of facial expressions was the correlation between their membership ratings in each category, calculated across the categories.³ The resulting similarity matrix was analyzed by the Guttman-Lingoes multidimensional scaling procedure, smallest space analysis-I (SSA-I) (Lingoes, 1965). The resulting structure is shown in Figure 2.

Was the obtained structure similar to that predicted by our circular model? To answer this question, the 14 facial expressions were scaled on the dimensions of pleasure-displeasure and arousal-sleepi-

FIGURE 2
Intercategory structure: Multidimensional scaling of similarity defined as the correlations between degree-of-membership ratings.



3. We also calculated a profile similarity measure, the sum of absolute values of differences between mean ratings. The resulting similarity matrix, analyzed by SSA-I, yielded a spatial representation nearly identical to that shown in Figure 2.

ness. Separate groups of subjects rated the emotional message of each of the 14 photographs on Mehrabian and Russell's (1974) pleasure and arousal scales. This corresponds to placing them within the two-dimensional circular structure proposed by Russell (1980). The pleasure scores correlated .93 with the scale values on the horizontal dimension of Figure 2. The arousal scores correlated .82 with the scale values of its vertical dimension. Moreover, the order of emotions around the perimeter of the two spaces corresponded exactly. In short, these results demonstrate that internal structure of emotion categories is closely tied to intercategory structure.

Discussion

If emotion categories fit the classical view, then we should have found a clear division between members and nonmembers of each category. Error in the data might not have allowed all 0's and 6's, but curves such as those in Figure 1 should have approximated step functions: nonmembers as close to the floor as random error allowed, members as close to the ceiling as random error allowed. Instead, membership in every category was graded. On the other hand, the task given to the subjects all but demanded that they produce graded responses. The more important point of this study is therefore not gradedness per se; it is that the gradedness was reliable and systematic. Task demand cannot account for the agreement among subjects in their ratings. Task demand also cannot account for the systematic nature of the graded membership functions: Membership functions corresponded to (indeed, sufficed to reproduce) the ordering predicted from an intercategory structural model of emotion. This structural model (Russell, 1980) was originally proposed to account, not for judgments about facial expressions, but for the judged similarity among emotion concepts.

REPLICATION AND EXTENSION

STUDY 2: FUZZY BORDERS

If the boundaries of an emotion category are, as we argue, fuzzy, then there ought to be occasional disagreement over whether a particular facial expression is in or out of the category. Almost everyone can agree that the prototypical expressions are inside, and the very poor examples are outside, the category. But expressions with an intermediate degree of membership in the category may be difficult to place in or out.

Graded membership and the lack of a clear-cut boundary might be thought to amount to the same thing, but this appears not to be so. Armstrong *et al.* (1983) showed that such classically definable concepts as "even number" and "odd number" have a graded membership function. The numbers 2, 4, and 8 exemplify the concept of "even number" better than do 72, 510, and 398. Still, the border between even and odd numbers is clear-cut: Subjects can unequivocally adjudicate membership in "even number," "odd number," and other classically defined concepts, although they cannot do so for concepts such as "vehicle," "fruit," or "sport" (Fehr & Russell, 1984). We therefore take lack of a clear-cut boundary as a necessary feature in our definition of a fuzzy concept. Following the Fehr and Russell (1984) method, we take consensus versus disagreement in adjudication of membership as the criterion of a clear-cut versus a fuzzy boundary.

In most previous work on how individuals categorize facial expressions of emotion, subjects were asked to select *one* label from a *set*—a procedure that leaves open the question of whether any other labels in the set are applicable. Even allowing subjects to select more than one label leaves an asymmetry: The subjects' *active* selection of some labels as appropriate is taken as the measure of inclusion in those categories; but a *default*, an omission of a response, is taken as exclusion from a category. We chose a more symmetrical approach in this study by asking subjects to say either "yes" or "no" for each of a set of labels applied to each facial expression they were shown.

Method

Fifty undergraduates were shown the 14 photographs of facial expressions from Study 1, one at a time. The subject was asked to reply either "yes" or "no" to the question "Is this person X?" with X replaced by one of the emotion labels, such as "happy" or "sad," from the preceding study. The question was repeated with the same emotion word until all 14 photographs had been seen. This procedure was then repeated for another six category labels. In all, each subject responded to every photograph seven times.

Results and Discussion

The number of subjects responding "yes" for each word applied to each expression is shown in Table 2. There was consensus in some cases (i.e., figures close to 0% or 100%), but there were also various degrees of endorsement between 0% and 100%. And the degree of endorsement did not reflect simply random error: The probability of

TABLE 2
Number of Subjects Endorsing an Emotion Label for a Facial Expression

CATEGORY	PHOTOGRAPH OF FACIAL EXPRESSION														N
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
Excited	22	12	1	—	—	—	—	—	3	2	8	7	5	15	
Happy	24	25	24	24	15	5	1	—	—	—	—	—	—	8	
Glad	25	24	23	20	12	2	2	—	—	—	—	1	—	12	
Calm	8	21	24	24	25	23	20	8	1	4	—	—	3	8	
Sleepy	—	1	1	6	15	9	22	6	2	—	—	—	1	1	
Sad	—	—	—	—	6	6	13	25	16	2	1	2	2	1	
Unhappy	—	—	—	—	4	7	11	25	23	19	18	17	4	1	
Miserable	—	—	—	—	4	5	9	25	23	12	12	11	1	—	
Disgusted	—	1	—	—	1	12	7	4	5	23	14	4	1	3	
Mad	—	—	1	—	1	3	2	3	7	12	24	11	2	1	
Angry	—	—	—	—	1	3	3	4	4	19	24	9	2	—	
Scared	—	—	—	—	—	1	1	5	15	—	3	23	19	5	
Afraid	—	—	—	—	—	—	—	5	9	1	3	20	13	4	
Surprised	19	4	1	—	—	—	—	—	2	1	4	14	25	25	

Note. Maximum number of subjects is 25.

endorsement correlated from .94 to .99 with the degree-of-membership ratings from Study 1.

The data from this study were also used to generate a representation of intercategory structure. Pairwise similarity was taken as the correlation between proportions of agreement and submitted to the Guttman-Lingoes SSA-I multidimensional scaling procedure. The result was nearly identical to that obtained in Study 1: The horizontal and vertical axes each correlated .99 with the corresponding axes of Figure 2 and correlated, not surprisingly, .95 and .84 with pleasure and arousal scores obtained in Study 1 for the 14 expressions.

These results show that graded membership of natural language categories of emotion can be demonstrated even when subjects are not asked to rate the degree of membership—indeed, when the task demands that they give an either-or answer. These results also show that one and the same study can replicate the seemingly conflicting empirical results from all three camps concerned with perception of emotion in facial expression: high or perfect agreement among observers for the prototype for each category; disagreement over the categorization of nonprototypes for that category; and a systematic intercategory structure revealing underlying dimensions.

STUDY 3: CHOICE OF EXEMPLARS⁴

One task commonly employed in research on facial expressions is called a recognition task (e.g., Izard, 1971). The subject is asked to choose which of various facial expressions belong to an emotion category ("Which person is angry?"). In Study 3, we asked whether a fuzzy-set approach could shed light on subjects' responses in this task. Subjects were shown an array of facial expressions and asked to find three exemplars for each category.

Method

Thirty-eight undergraduates were shown nine of the photographs used in the preceding two studies. There was one photograph for each of the following categories: A, excitement; B, happiness; C, calmness; D, sleepiness; E, sadness; F, disgust; G, anger; H, fear; and I, surprise. There was also a 10th photograph, which was a neutral expression taken from Ekman and Friesen (1976). The 10 were

4. This study was reported in another paper (Bullock & Russell, 1984) in order to compare the responses of children with those obtained from adults on the same task.

of different female faces, sex being held constant to remove it as a factor influencing subjects' choices. Of the 14 emotion categories studied in the two preceding studies, 12 were included here; "unhappy" and "glad" were omitted.

The experimenter showed each subject the 10 photos, spreading them out in a random array. The experimenter asked "Which person is X? Can you show me the X person?" with X replaced by the label for one of the emotion categories. After the subject chose a picture, the experimenter removed it from the array and repeated the question with the same emotion word. The second picture chosen was removed, and the question was repeated a third time. In this procedure, the three choices for each word constituted one trial. Each subject was given 15 trials in all—one for each of the 12 emotion words reported here, plus 3 for other words not of interest here. Trials were in a different random order for each subject, and between each trial the set of photographs was shuffled and set out in a new random order.

Results and Discussion

The number of subjects who selected each photograph is given in Table 3. To illustrate the results, Figure 3 shows a histogram for the category "angry." Nine photographs (the neutral being omitted) are placed along the horizontal axis in the order derived from our circular structural model. The vertical axis shows the percentage of subjects who chose each photograph as an exemplar of "angry." Ordering the photographs in this way showed that subjects' choices were clustered in a systematic fashion. On their first choice, almost 90% chose photograph G, which is Ekman and Friesen's (1976) prototypical expression of anger and the expression that received the highest degree-of-membership rating for "angry" in Study 1. But those who chose otherwise (who made "errors" in the classical view) were clearly not choosing at random: they either picked face F or face H, which are the two facial expressions adjacent to G according to our model and the two with the next highest degrees of membership in the category "angry."

This conclusion is reinforced by turning to second and third choices, which also tended to be of the two expressions adjacent to face G. Those few subjects who chose otherwise tended to pick expressions adjacent to these.

These results thus demonstrate the influence of graded membership on choices in a "recognition" task. The prototype was typically chosen first. Once the prototype was removed from the set, second

TABLE 3
Number of Subjects Choosing a Facial Expression as Exemplar of a Category

WORD	CHOICE	PHOTOGRAPH OF FACIAL EXPRESSION								
		A	B	C	D	E	F	G	H	I
Excited	1	35	2	—	—	—	—	1	—	—
	2	2	23	—	—	—	—	6	3	4
	3	—	4	4	—	—	2	11	8	7
Happy	1	22	16	—	—	—	—	—	—	—
	2	14	22	—	1	—	1	—	—	—
	3	—	—	35	1	—	—	—	—	—
Calm	1	—	4	18	5	—	—	—	—	—
	2	1	—	17	10	—	—	—	—	—
	3	—	5	3	18	—	—	—	—	—
Sleepy	1	—	—	1	33	3	—	1	—	—
	2	—	—	20	4	9	—	1	—	—
	3	—	—	12	1	11	1	—	—	—
Sad	1	—	—	—	2	36	—	—	—	—
	2	—	—	—	29	2	1	1	—	—
	3	—	—	1	5	—	1	2	2	3
Miserable	1	—	—	—	—	33	2	3	—	—
	2	—	—	—	16	2	9	8	2	—
	3	—	—	—	2	2	14	12	2	2
Disgusted	1	—	—	—	2	1	32	1	1	—
	2	—	—	1	10	7	4	11	2	2
	3	—	—	1	6	10	—	11	3	—
Mad	1	—	—	—	—	1	1	36	—	—
	2	—	—	—	—	—	22	1	7	—
	3	—	—	—	1	3	13	1	9	—
Angry	1	—	—	—	—	—	2	34	2	—
	2	1	—	—	1	1	26	2	4	—
	3	—	—	—	—	6	3	1	12	1
Scared	1	—	—	—	—	1	—	—	31	6
	2	—	—	—	—	4	—	1	7	26
	3	1	—	—	—	13	1	12	—	4

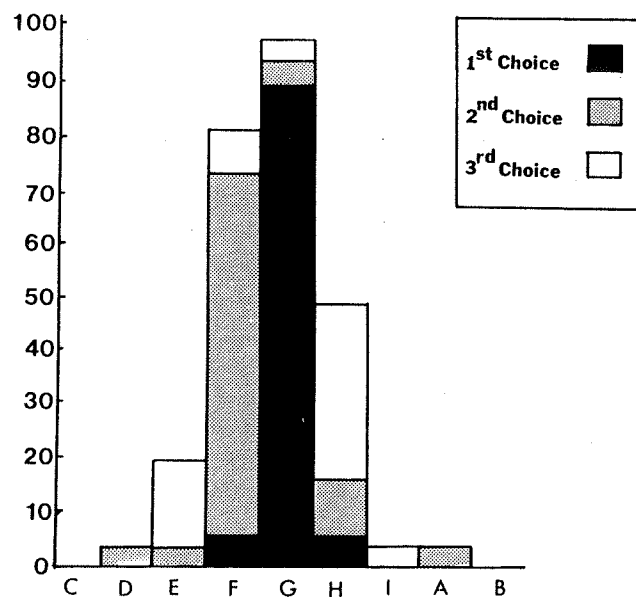
(continued)

TABLE 3 (Continued)

WORD	CHOICE	PHOTOGRAPH OF FACIAL EXPRESSION								
		A	B	C	D	E	F	G	H	I
Afraid	1	—	—	—	—	2	—	—	30	6
	2	1	—	—	—	7	1	3	5	21
	3	1	1	—	—	13	6	8	—	9
Surprised	1	10	—	—	—	—	—	—	1	27
	2	20	1	—	—	—	—	—	7	10
	3	5	10	—	—	—	2	3	17	1

Note. $n=38$ for each choice. Subjects chose from the 9 photographs listed plus a 10th, neutral expression. When the numbers in a row sum to less than 38, the remaining subjects selected the neutral expression.

FIGURE 3
Choice of expressions as exemplars of anger.



and third choices were not random—subjects chose expressions with the next highest degree of membership in the category. Degree-of-membership scores from Study 1 thus ought to predict the subjects' choices observed in this study. Correlations between the degree-of-membership ratings in a category and the probability of being chosen (either first, second, or third) as an exemplar of that category ranged from .73 to .99.

STUDY 4: EXPRESSIONS OF UNIVERSAL EMOTIONS IN A NEW GUINEA VILLAGE

The preceding three studies relied heavily on one set of posed facial expressions. Would we obtain similar results with expressions of genuine feelings? To answer this question properly would require good photographs of spontaneous expressions, plus some assurance that the expressions chosen were as pure examples of what have been considered basic emotions as could be obtained. Fortunately, a book published by Paul Ekman (1980) provided just what we needed.

Ekman's (1980) book, *The Face of Man: Expressions of Universal Emotions in a New Guinea Village*, contains over 50 finely reproduced photographs taken by Ekman of spontaneous facial expressions of people living in a Stone Age village. As Ekman pointed out, "there is an enormous advantage to being with people who were not camera-shy. They did not know what a camera did so they were not self-conscious about it" (p. 11). Ekman selected photographs that, in his judgment, best represent basic emotions:

[T]he photographs are grouped to show the six emotions we studied, all of which we found to be universal in appearance: happiness, fear, surprise, anger, disgust, and sadness. There are also pictures of perplexity, interest, and embarrassment, which we believe to be universal but which we did not specifically study. Since these pictures show universal facial expressions, the message conveyed by each face will usually be quite obvious. . . . All of the photographs were made to document what I learned. They simply tell the story of what I found. . . . [T]he best argument for universality is made by the faces of the New Guineans." (pp. 11-12)

In this study, we asked subjects to examine each photograph and to rate how accurately a word (such as "happy") described the feeling of the person shown. We used a rating of accuracy because we thought it would be an alternative index of degree of membership that would be easier for subjects to use.

Method

Subjects were 50 undergraduates of the University of British Columbia who received course credit for their participation.

Six photographs were selected from Ekman's (1980) book that appeared to us to be the clearest expressions of six different emotions: surprise, fear, anger, disgust, sadness, and interest. For each photograph, the plate number and label given by Ekman (1980) are listed in Table 4.

To ensure maximum fidelity to Ekman's original photograph, subjects were shown the book (captions were masked). The subject was asked to examine the particular face in the book and to "look for the feeling that is expressed there." A judgment scale was provided listing the eight emotion-descriptive categories given in Table 4. The subject was asked to rate how accurately each word described the feeling expressed. Accuracy was judged on a scale from 1 ("extremely inaccurate") to 8 ("extremely accurate").

The experimenter then turned to another page in the book and asked for a rating of another facial expression, repeating this procedure until all six expressions had been rated.

Results and Discussion

Table 4 gives the mean judged accuracy of each word as a descriptor for each face. To estimate the reliability of these values, a correlation between the ratings from two randomly created subsamples were calculated for each face. The results, given in the bottom row, indicated that the ratings were highly reliable.

Table 4 shows the pattern of results now familiar from the first three studies. Consider the category "angry." Plate 17, presented by Ekman as a facial expression of anger, received the highest average rating. However, plates 14, 16, and 10 were also reasonably good examples of angry, with ratings only moderately lower than those for plate 17 and higher than plates 22 or 7. (Individual ratings showed a similar pattern of intermediate values. In other words, the in-between ratings of accuracy were not an artifact resulting from averaging very high and very low values.) The results for the other categories listed showed a similar pattern: The plate given by Ekman to exemplify a particular emotion usually emerged as the prototype for that category, but other plates exemplified the category to varying degrees.

One curious finding is the high ratings across all six plates for the category "calm." "Calm" was uniformly rated a more accurate descriptor than its place in our circular structural model would have

TABLE 4
Mean Rating on Accuracy of Category as Descriptor of a Facial Expression

CATEGORY	FACIAL EXPRESSION					
	22 [INTEREST]	14 [SADNESS]	16 [DISGUST]	17 [ANGER]	10 [FEAR]	7 [SURPRISE]
Excited	4.5	3.0	4.3	3.4	3.5	3.6
Happy	5.1	2.1	2.3	1.7	2.2	3.8
Calm	5.2	4.0	3.1	4.4	3.6	5.0
Sleepy	2.5	3.1	2.2	3.0	3.3	2.4
Bored	3.2	3.3	3.1	4.4	3.4	3.5
Sad	3.1	5.9	3.2	5.3	4.5	4.0
Disgusted	3.1	4.6	6.8	5.6	4.4	3.0
Angry	2.9	4.8	5.7	6.0	5.2	2.9
Afraid	3.3	5.8	3.7	4.9	5.8	4.3
Surprised	4.8	4.2	3.6	3.0	4.7	6.6
Reliability	.89	.95	.96	.95	.93	.85

Note. Ratings were made on a scale from 1 ("extremely inaccurate") to 8 ("extremely accurate"). The numbers for facial expressions refer to plates in Ekman's (1980) book. Label in brackets is that provided by Ekman.

predicted. Looking through other photographs in Ekman's book, we were struck by an apparent calmness in the people portrayed. Of course we have insufficient data to go on, but perhaps this high rating on "calm" reflects a constant difference between these New Guinean villagers and our North American raters.

DISCUSSION OF THE FOUR STUDIES

These four studies were not designed to test the thesis that basic-level categories of emotion are fuzzy. Rather, they were designed to see whether subjects' categorization of facial expressions can be more easily understood when this is assumed.

In four different tasks, subjects failed to divide facial expressions into mutually exclusive, all-or-none categories. Instead, they treated facial expressions as though they could be ordered according to the degree to which they were members of various categories. They did so for all the facial expressions and emotion categories we examined. Subjects produced approximately the same graded responses whether implicitly asked to do so (Studies 1 and 4) or not (Studies 2 and 3). Subjects failed to treat the categories specified by such words as "angry," "afraid," and "happy" as mutually exclusive: In all four studies, the same expressions were found, to varying degrees, in different categories. And subjects failed to agree with one another about the borders between categories. Ordering facial expressions according to the degree to which they are members of emotion categories thus predicted the extent of disagreement found in a forced-choice task, as well as the pattern of "correct" and "incorrect" choices in a recognition task.

Disagreement found in the past among subjects categorizing the same facial expression was often attributed to "ambiguity" in the facial expression. It is thus perhaps worth emphasizing that disagreement occurred here even with photographs carefully chosen as the most likely to represent one and only one category. More important, the cases of disagreement seen here were not due to the amount of information available in the particular expression: The *same* photograph elicited consensus on one label and disagreement on another label. At least in the cases studied here, the basis of disagreement is not perceptual, but semantic. Further support for this contention comes from the multidimensional scaling results of Study 2: A structural model proposed on the basis of similarity of word meaning predicted the pattern and degree of disagreement about facial expressions. Of course, the information on which categorization is based is in the

face, and in this sense, the present interpretation complements rather than contradicts efforts to explain disagreements in terms of physical features.⁵

In short, although the evidence presented cannot prove that emotion categories are overlapping fuzzy sets rather than mutually exclusive proper sets, our evidence is consistent with that thesis. In other domains, when categorization of actual objects or events has shown such a pattern, more definitive semantic tests of fuzziness have been found to be positive (Smith & Medin, 1981). We therefore feel justified in resuming our attempt to draw out the implications of assuming that emotion categories are fuzzy. We next elaborate on the meaning of a facial expression, the role of the dimensions of pleasure and arousal, and how it is possible for information about context to disambiguate the meaning of a facial expression.

THE PRIMARY MEANING ATTRIBUTED TO A FACIAL EXPRESSION

So far, we have focused on the emotion category. We now turn to the facial expression, asking if we can characterize the *primary* meaning that a person reads from a particular expression. (By "primary meaning," we mean the message attributed to the facial expression alone:

5. The word "disagreement" requires a cautionary note. If, as we believe, some of the traditional assumptions about emotion categories are incorrect, then methods based on these assumptions may produce artifactual results. To assess how a subject categorizes facial expressions, researchers have, for example, asked the subject to choose for the expression one label from a list of six "basic" emotions. Any such method that requires either-or answers or that presets the number of choices presupposes that emotion categories are nonoverlapping proper sets. Investigations based on such methods appeared to show persistent "disagreements" and, presumably, "inaccuracy." Subjects' labeling of facial expressions appeared fraught with error. Only when photographs were highly selected and subjects' choices limited to a half-dozen "basic" emotions did disagreement disappear—and even then not entirely (Ekman & Friesen, 1976; Tomkins & McCarter, 1964). Thus, Tomkins and McCarter (1964) wrote at great length about why disagreement occurs. From our perspective, disagreement may be largely an artifact of the methods commonly employed.

If subjects want to reply with an answer halfway between "yes" and "no," then forcing them to choose will result in the appearance of disagreement. If subjects want to pick two categories, then forcing them to choose one will result in the appearance of disagreement. Given the opportunity to select various categories and to give answers that reflect degrees of classification intermediate between "yes" and "no," subjects do so—and are highly consistent with one another. The most striking feature of our four studies is the amount of agreement among subjects. From a prototype perspective, their behavior is not erroneous or anomalous, but orderly and systematic.

Even when we believe that a smile is feigned or hides grief, we recognize the intended meaning of the smile itself.)

The primary meaning attributed to a facial expression has generally been assumed to be *one* emotion category (or, in the case of "blends," two categories). This assumption is implicit when an expression is labeled with a single emotion term: an "anger expression" (or an "anger-fear blend"). Our results suggest that a characterization of primary emotional meaning as a single category is but a first approximation. Even a highly selected prototypical expression is a member, to varying degrees, of several emotion categories. For more natural facial expressions, a single category is likely to be even less appropriate.

To be specific, let us return to the empirical results reported in Table 1. Rather than look at an emotion category (represented as a row), let us focus on one facial expression (represented as a column). Consider photograph N, the last column of Table 1. This photograph is given in Ekman and Friesen's (1976) collection of posed expressions to represent surprise. And, indeed, its highest membership rating in Study 1 was for "surprised." If subjects were forced to pick one label for N, "surprised" would clearly be the best choice (and "surprised" is what in Study 2 they tended to choose). But the meaning of expression N, as revealed by looking over the full range of ratings in the column, is not well captured by one or even several categories. To illustrate, Figure 4 shows the degree of membership of photograph N in each of 12 categories, with the categories ordered according to our structural model.

Also shown in Figure 4 is a plot for photograph M, another prototypical expression of surprise from Ekman and Friesen's (1976) collection. This plot shows that the basic result found for N was replicated with M: Although "surprised" was the single best descriptor, nearby categories applied to varying degrees. More importantly, M and N differed in a way not captured by classifying both as expressions of surprise. The curve for M in Figure 4 is slightly to the left of that for N. That is, M held a moderate degree of membership in "afraid" and a slight degree of membership in most negative categories. N, on the other hand, was a poorer member of "afraid" and other negative categories, but was a better member of "excited," "happy," and "glad," the positive categories. We have replicated this difference elsewhere, which thus suggests that even the best exemplars within a category can reliably vary in subtle ways in the meaning they convey. Again, a single label fails to capture these differences.

These considerations suggest, to us, that the primary meaning attributed to a facial expression involves a two-step process. A facial expression is initially and automatically perceived in terms of degree of

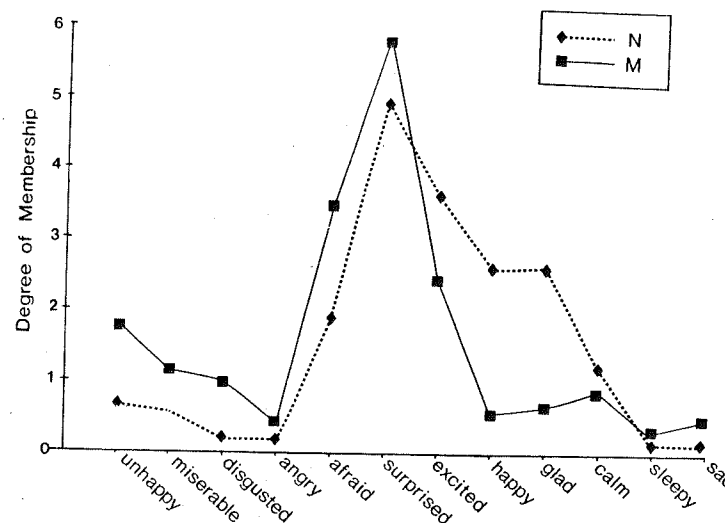


FIGURE 4

Degree of membership in 12 categories of emotion for two prototypical surprise expressions from Ekman and Friesen's (1976) collection. (N=M01-14; M=PF1-16).

pleasure and degree of arousal. Phenomenologically, these dimensions combine into an integral, unitary perception: The face expresses, say, pleasure-and-arousal, or displeasure-and-low-arousal. This step can be graphically represented as locating the face within the spatial structural model depicted in Figure 2. (Notice that the location of the two "surprise" pictures, M and N, in Figure 2 reflects the subtle differences in their categorization seen in Figure 4.) Next, to articulate the meaning of the expression, it can be associated with labels for one or more emotion categories. In some circumstances, a single label is useful, and the closest one can be picked. In other circumstances, a richer description can be achieved by specifying the expression's degree of membership in (closeness to) several categories. More generally, the applicability to face of any category label is inversely proportional to the distance between the face and the label.

If our line of reasoning is correct, how does one person ever perceive another as, say, afraid and only afraid? To answer this question, we must consider the more typical case in which a person categorizes the emotion of another on the basis not only of facial expression, but of many sources of information.

PERCEPTION OF EMOTION IN OTHERS

Consider the prototypical case of classifying a man as afraid: A danger appears. He gasps and stares at it. His eyes bulge, his jaw drops, his heart pounds, his hands tremble. He is frozen with panic. His eyes dart about. Suddenly, he turns and flees.

When we say that someone is afraid, we combine in one word a series of consecutive events dependent on one another. To know the meaning of the word "fear" is to know that sequence. It is to know a "script" (Abelson, 1981) in which events unfold in a certain order. The concept—the script—specifies the prototypical elements, their order, and their causal connections: context, appraisal of the context, physiological reactions, feelings, vocal expressions, facial expressions, actions, and consequences.

In perceiving the emotion of another, a person implicitly constructs from the available information a plausible scenario, including inferences about unobserved aspects of the other's current state and predictions about its future course. The scenario constructed may resemble various emotion scripts to varying degrees. The generalized script both aids in the construction of the particular scenario and constitutes the ideal against which the particular is compared. The degree of membership in an emotion category is the degree to which the particular scenario constructed resembles the idealized script.

The notion of script helps explain how it is possible for the situation to "disambiguate" a facial expression of emotion. According to our view, contextual information is part of the concept of fear. Therefore, the more the context is a prototypical fearful situation, the more applicable is the label of "fear." The same could be said about any element in the script: The person's behavior, signs of physiological activity, and introspected feelings all "disambiguate" the emotion by increasing (or decreasing) resemblance to the script. But neither feelings nor facial expressions nor situations are *defining* features. Thus, we could equally well say that the facial expression disambiguates information from context as vice versa.

Of course, fear can be said to exist even when the fear script is not matched perfectly. Just as the category of "fear" allows a range of different facial expressions, it allows a range of manifestations in each element of the script. For example, we know that the man described above might inhibit some of the physiological or expressive signs of fear. He might remain frozen rather than flee. The situation might not actually be dangerous. He might not even believe the situation is dangerous, as in the case of phobics who admit that their fear is unreasonable. Of course, if enough events in the sequence are absent or

altered or out of order, we are unlikely to call it fear. But the border is fuzzy, and there exist cases in which so few of these events occur that one is unsure whether the label "fear" is appropriate or not.⁶

CONCLUDING REMARKS

Those who study emotion often divide themselves into camps, on the assumption that their different positions conflict. Hierarchical models are said to conflict with circumplex models, categories with dimensions, and so on. The picture emerging from our study is that at least some of these notions, when properly understood, are compatible. For example, we have argued that the domain of emotion concepts is structured in at least three senses, senses sometimes thought to be in conflict. First, we said that emotion categories can be arranged in a *hierarchical structure*—meaning that "emotion" includes "fear," "anger," "love," and an indefinite number of others, and that at least some of these middle-level categories are further subdivided. Second, each emotion category possesses an *internal structure*—meaning that members vary in the degree to which they are members. And third, the full set of middle-level emotion categories possesses an *intercategory structure*—meaning that each category is related to every other category in the domain in a specifiable manner. Intercategory structure is related to internal structure in that it is a representation of the similarities between internal membership functions. These three types of structure are in no more conflict than are wave and particle theories of light: Each highlights different properties of the way in which people conceptualize emotion.

Anyone who says that the domain of emotion is structured in these ways, or that concepts such as "emotion," "fear," or "anger" are fuzzy, or that they are scripts, risks being misunderstood. We have not said that an *event* of fear is fuzzy, scripted, or structured. Nor have we said that the "true" meaning of a facial expression is as shown in Figure 4. We have been concerned with the mental representation of, rather than the event of, emotion. The properties of concepts must be distinguished from the properties of events. Perhaps some of the properties of emotion categories in some sense reflect

6. We have little evidence to go on, and the hypotheses sketched here probably oversimplify matters. We are assuming that, just as a facial expression has a primary meaning, each element of the script has a primary meaning. We are also assuming that each observed element simply adds to or subtracts from an overall resemblance to the script. How reasonable these assumptions are remains to be seen.

properties of actual emotion events. Or perhaps some of the properties of emotion categories reflect principles of mental processing, rather than emotions uniquely. To claim that natural language concepts of emotion are fuzzy is therefore not to claim that any actual event denoted by the word "emotion" is fuzzy.

We also hope that no one will take us as claiming that a scientific taxonomy of emotion events should consist of overlapping fuzzy categories. Natural language often provides fuzzy concepts where precise ones are possible. For example, the number of elements in a set can be precisely specified (one, two, three, etc.), but English also allows us to specify number in an imprecise way ("a few," "some," "many"). Some languages (because the exact number is unknown, "some" is a handy concept) lack a system of numbers other than what might be translated as "one," "two," and "more than two." Speakers of such languages should resist the temptation to assume that their handy folk taxonomy of numbers is the best taxonomy possible. Now consider the rather vague natural language concepts we have for types of facial expression: "smile," "smirk," "grin," "frown," "grimace," "sneer," "pout," and so on. Ekman and Friesen (1978) developed a descriptive taxonomy for facial actions that shows that facial expressions are capable of much more precise description. "Happiness," "anger," "fear," and the rest are concepts we inherit from our culture to distinguish types of events. Fuzziness is a fact about the concepts, not about the events. A psychological account on how ordinary people conceptualize emotion events must recognize that fuzziness. Nevertheless, fuzziness must not be taken as a limit on how we as scientists can or should describe those events.

REFERENCES

- Abelson, R. P. (1981). Psychological status of script concept. *American Psychologist*, 36, 715-729.
- Abelson, R. P., & Sermat, V. (1962). Multidimensional scaling of facial expressions. *Journal of Experimental Psychology*, 63, 546-554.
- Armstrong, S. L., Gleitman, H., & Gleitman, L. R. (1983). What some concepts might not be. *Cognition*, 13, 263-308.
- Brown, R. (1980). *Natural categories and basic objects in the domain of persons*. The Katz-Newcomb Lecture, University of Michigan, Ann Arbor, MI.
- Bruner, J. S., & Tagiuri, R. (1954). The perception of people. In G. Lindzey (Ed.), *Handbook of social psychology* (Vol. 2, pp. 634-654). Reading, MA: Addison-Wesley.
- Bullock, M., & Russell, J. A. (1984). Preschool children's interpretation of facial expressions of emotion. *International Journal of Behavioral Development*, 7, 193-214.

- Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review*, 90, 105-126.
- Cantor, N., & Mischel, W. (1977). Traits as prototypes: Effects on recognition memory. *Journal of Personality and Social Psychology*, 35, 38-48.
- Cantor, N., & Mischel, W. (1979). Prototypes in person perception. *Advances in Experimental Social Psychology*, 12, 38-48.
- Cantor, N., Smith, E. E., French, R. D., & Mezzich, J. (1980). Psychiatric diagnosis as prototype categorization. *Journal of Abnormal Psychology*, 89, 181-193.
- Dahlgren, K. (1985). The cognitive structure of social categories. *Cognitive Science*, 9, 379-398.
- Daly, E. M., Lancee, W., & Polivy, J. (1983). A conical model for the taxonomy of emotional experience. *Journal of Personality and Social Psychology*, 45, 443-457.
- Darwin, C. (1965). *The expression of the emotions in man and animals*. Chicago: University of Chicago Press. (Original work published 1872)
- Ekman, P. (1972). Universals and cultural differences in facial expressions of emotions. In J. K. Cole (Ed.), *Nebraska Symposium on Motivation*, 1971 (Vol. 19, pp. 207-283). Lincoln: University of Nebraska Press.
- Ekman, P. (1980). *The face of man: Expressions of universal emotions in a New Guinea village*. New York: Garland STPM Press.
- Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., & Friesen, W. V. (1978). *The Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., Friesen, W. V., & Ellsworth, P. (1982). Does the face provide accurate information? In P. Ekman (Ed.), *Emotion in the human face* (2nd ed., pp. 56-97). Cambridge, England: Cambridge University Press.
- Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in the facial displays of emotion. *Science*, 164, 86-88.
- Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, 113, 464-486.
- Fehr, B., Russell, J. A., & Ward, L. M. (1982). Prototypicality of emotions: A reaction time study. *Bulletin of the Psychonomic Society*, 20, 253-254.
- Fernberger, S. W. (1928). False suggestion and the Piderit model. *American Journal of Psychology*, 40, 562-568.
- Hampson, S. E. (1982). *The construction of personality*. London: Routledge & Kegan Paul.
- Horowitz, L. M., Wright, J. C., Lowenstein, E., & Parad, H. W. (1981). The prototype as a construct in abnormal psychology: I. A method for deriving prototypes. *Journal of Abnormal Psychology*, 90, 568-574.
- Hunt, W. A. (1941). Recent developments in the field of emotion. *Psychological Bulletin*, 38, 249-276.
- Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.
- James, W. (1929). *The varieties of religious experience: A study in human nature*. New York: Longmans, Green. (Original work published 1902)
- Jones, G. V. (1982). Stacks not fuzzy sets: An ordinal basis for prototype theory of concepts. *Cognition*, 12, 281-290.
- Labov, W. (1973). The boundaries of words and their meanings. In C. J. N. Barley & R. W. Shuy (Eds.), *New ways of analyzing variation in English* (pp. 340-373). Washington, DC: Georgetown University Press.
- Lakoff, G. (1973). Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*, 2, 458-508.
- Landis, C. (1924). Studies of emotional reactions: II. General behavior and facial expres-

- sion. *Journal of Comparative Psychology*, 4, 447-509.
- Landis, C. (1929). The interpretation of facial expression in emotion. *Journal of General Psychology*, 2, 59-72.
- Langer, S. K. (1980). *Philosophy in a new key*. Cambridge, MA: Harvard University Press. (Original work published 1942)
- Lingoes, J. C. (1965). An IBM-7090 program for Guttman-Lingoes smallest space analysis-I. *Behavioral Science*, 10, 183-184.
- McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory and Cognition*, 6, 462-472.
- Mehrabian, A. (1972). *Nonverbal communication*. Chicago: Aldine-Atherton.
- Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge, MA: MIT Press.
- Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, 32, 89-115.
- Neisser, U. (1979). The concept of intelligence. *Intelligence*, 3, 217-227.
- Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9, 35-38.
- Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, 35, 677-688.
- Rosch, E. (1973). Natural categories. *Cognitive Psychology*, 4, 328-350.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192-233.
- Rosch, E. (1977). Human categorization. In N. Warren (Ed.), *Studies in cross-cultural psychology* (pp. 1-49). London: Academic Press.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161-1178.
- Russell, J. A. (1984). *Fuzzy concepts and the experience of emotion*. Unpublished.
- 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.
- Russell, J. A., & Bullock, M. (1986). On the dimensions preschoolers use to interpret facial expressions of emotion. *Developmental Psychology*, 22, 97-102.
- Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology*, 44, 229-237.
- Schneider, D. J., Hastorf, A. H., & Ellsworth, P. C. (1979). *Person perception*. Reading, MA: Addison-Wesley.
- Sherman, M. (1927a). The differentiation of emotion responses in infants: I. Judgments of emotional responses from motion pictures views and from actual observation. *Journal of Comparative Psychology*, 7, 265-284.
- Sherman, M. (1927b). The differentiation of emotional responses in infants: II. The ability of observers to judge the emotional characteristics of the crying of infants and the voice of an adult. *Journal of Comparative Psychology*, 7, 335-351.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- Spignesi, A., & Shor, R. (1981). The judgment of emotion from facial expressions, contexts and their combination. *Journal of General Psychology*, 104, 41-58.
- Tomkins, S. S. (1962-1963). *Affect, imagery, consciousness* (2 vols.). New York: Springer.
- 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.
- Wagner, H. L., MacDonald, C. J., & Manstead, A. S. R. (1986). Communication of

- individual emotions by spontaneous facial expression. *Journal of Personality and Social Psychology*, 50, 737-743.
- Wiggins, J. S. (1980). Circumplex models of interpersonal behavior. In J. Wheeler (Ed.), *Review of personality and social psychology* (pp. 265-294). Beverly Hills, CA: Sage.
- Wittgenstein, L. (1953). *Philosophical investigations*. New York: Macmillan.
- Woodworth, R. S. (1938). *Experimental psychology*. New York: Holt.
- Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology*. New York: Holt, Rinehart, & Winston.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8, 338-353.
- Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology*, 43, 111-122.