Conceptual and Methodological Issues in the Judgment of Facial Expressions of Emotion[1]

Erika L. Rosenberg[2] and Paul Ekman
University of California, San Francisco

In two studies, subjects judged a set of facial expressions of emotion by either providing labels of their own choice to describe the stimuli (free-choice condition), choosing a label from a list of emotion words, or choosing a story from a list of emotion stories (fixed-choice conditions). In the free-choice condition, levels of agreement between subjects on the predicted emotion categories for six basic emotions were significantly greater than chance levels, and comparable to those shown in fixed-choice studies. As predicted, there was little to no agreement on a verbal label for contempt. Agreement on contempt was greatly improved when subjects were allowed to identify the expression in terms of an antecedent event for that emotion rather than in terms of a single verbal label, a finding that could not be attributed to the methodological artifact of exclusion in a fixed-choice paradigm. These findings support two conclusions: (1) that the labels used in fixed-choice paradigms accurately reflect the verbal categories people use when free labeling facial expressions of emotion, and (2) that lexically ambiguous emotions, such as contempt, are understood in terms of their situational meanings.

Over the past 25 years numerous studies of literate and preliterate cultures have shown that anger, disgust, fear, happiness, sadness, and surprise are universally recognized (for reviews see Ekman, 1989; Ekman, Friesen, & Ellsworth, 1972; Ekman & Oster, 1979; Izard, 1971, 1977). This finding,

which may be one of the more stable and reproducible phenomena in psychology, has served as a foundation for evolutionary and biological theories of emotion (e.g., Ekman, 1977; Tooby & Cosmides, 1987) as well as a point of reference for cognitive theories that propose universality in the antecedents of emotion (e.g., Lazarus, 1991).

Some criticisms of the cross-cultural studies on the judgment of faces have been raised in recent years (Russell, 1994; Wierzbicka, 1986). These range from the bases for the selection of stimulus slides to the choice of response formats. We are particularly concerned with the criticism of the almost standard use of fixed-choice[3] paradigms. Most of the studies reporting evidence in support of universals in facial expressions of emotion have employed a fixed-choice paradigm, in which subjects select from a list of emotion labels the one word that best describes the facial expression of a person shown in a photograph or slide. In most studies, consensus among subjects has been measured in terms of the percentage who selected each of the various labels for an expression. Recently, Russell (1994) questioned the validity of the universality findings by criticizing the use of the fixed-choice method. He said that this response format does not indicate which labels people would attribute to various expressions if they were given the opportunity to provide free responses, and that people might identify quite different emotions if they were not constrained by the fixed-choice set of categories. Furthermore, Russell suggested that the fixed-choice method pressures subjects to say that they see but one emotion in each expression when they may in fact see several emotions or none at all.

Although the results from the few facial judgment studies that have employed other methods are consistent with those from fixed-choice studies, they do not decisively address all of the issues raised by Russell. The problem of exclusivity of choice imparted by fixed-choice formats has been addressed by allowing subjects to make quantitative ratings on six emotion scales (c.f., Ekman et al., 1987). Subjects consistently rated at the highest intensity the same labels that are chosen most often to describe the same expressions when the fixed-choice method is used. The quantitative rating method, however, does not provide information on which words are the best labels for certain facial expressions of emotion. This problem can only be addressed through the use of a free-response format, wherein subjects provide labels of their own choice to describe facial expressions. Izard (1971) employed a free-response format and reported results that were consistent with those from fixed-choice studies. He did not, however, report

---

[3]We use the word *fixed* to refer to judgment conditions in which the response options are constrained. We believe this phraseology is more operational than *forced-choice*, which is what constrained choice judgment conditions have been called traditionally.

how he categorized the individual words that subjects gave in the free-re-
sponse task, which makes it difficult to evaluate whether it was the actual
free responses or Izard's categorizations of them that were consistent with
previous fixed-choice studies.

The present research had two major purposes. The first was to address
the methodological criticisms raised by Russell (1994) and others (Wierzbicka,
1986) regarding the fixed-choice paradigm. We sought to determine whether
a methodology in which subjects were allowed to provide any label to describe
a facial expression would yield evidence of agreement on the same or similar
emotion labels as those that have been used in fixed-choice studies of the
past. The second purpose was to deal with a conceptual problem concerning
the judgment of the facial expression of contempt, an emotion that has been
the source of controversy in the facial judgement literature.

The Special Case of Contempt

While at least three cross-cultural studies have produced evidence that
observers recognize a unilateral lip tightening as an expression of contempt
(Ekman & Friesen, 1986; Ekman & Heider, 1988; Matsumoto, 1992), the
universality, morphology, and discovery of the contempt expression have
been debated in the emotion literature in recent years (Ekman, O'Sullivan,
Matsumoto, 1991a, 1991b; Izard & Haynes, 1988; Russell, 1991a, 1991b,
1991c). The primary evidence for contempt's universal status comes from
the work of Ekman and his colleagues (Ekman & Friesen, 1986; Ekman
& Heider, 1988), although Izard (1971) obtained some evidence for rec-
ognition of contempt using a slightly different form of the expression. The
procedures used in those studies were typically the same as the fixed-choice
methods from other judgment studies: Subjects viewed a series of different
facial expressions and chose from a list of words which emotion the person
in slide was most likely to be feeling. On the basis of such research, the
agreement between subjects that a certain expression (usually unilateral lip
tightening) represented contempt was about 75%. This level of agreement
was significantly greater than would have been predicted by chance and
was obtained across cultures (Ekman & Heider, 1988; Matsumoto, 1992).
The agreement for contempt is comparable to but slightly lower than that
typically found for other emotions.

Much lower agreement levels for contempt have been reported by
other investigators. Ricci-Bitti, Brighetti, Garotti, & Boggi-Cavallo (1989)
found that Italian and American subjects were better at recognizing con-
tempt expressions posed by members of their own culture than they were
at recognizing contempt in members of the other culture. However, these

results are difficult to interpret, because it is not clear whether the same facial configurations were shown to both groups and whether the Italian translation of the word *contempt* had the same meaning as the word does in American English. Russell (1991a, 1991b, 1991c) also reported low agreement on the contempt expression of unilateral lip tightening. However, major methodological differences between Russell's studies and those of Ekman and his colleagues (Ekman & Friesen, 1986; Ekman & Heider, 1988) make it difficult to evaluate the differences in agreement levels for contempt obtained across these studies (see Ekman et al., 1991a, 1991b, for a detailed discussion).

The ambiguous findings on contempt may be due at least in part to the fact that the word *contempt* is not understood as well as other emotion labels, that it is used infrequently, and/or that many people are not able to distinguish its meaning from that of semantically similar emotion labels, such as anger and disgust (Ekman et al., 1991b). Our experience in previous research suggests that American subjects sometimes need to have the word defined for them. *Contempt* was not included in a large listing of over 2,000 English words considered suitable for the teaching of English as a foreign language (West, 1953), although anger, disgust, fear, happiness, sadness, and surprise were.

The problem of contempt raises a more general conceptual issue in the study of people's understanding of facial expression of emotion, which is whether single words are the best or only way to capture what an expression signals. According to Ekman (1989):

> Emotion terms can be thought of as a kind of shorthand, an abbreviated way to refer to a package of events and processes which comprise the phenomenon. Each emotion term . . .refers to a different set of organized integrated processes. They include antecedent events, the physiological and motor responses, the memories, thoughts, images and information processing, and the mobilization of efforts to cope with the source of emotion. All or any of these may be implied when someone say "he looks angry" (p. 159).

It is unlikely that facial expressions evolved to represent verbal labels or that the meaning of an expression is best captured by a particular emotion word (Ekman, 1994). Levy (1973, 1978, 1984) has reported that, in Tahitian culture, people show behavioral evidence of sadness in situations that typically evoke sadness in other cultures, even though they do not have a verbal label for this emotion. It is quite possible that people understand what the facial expression of "contempt"[4] implies in terms of the antece-

---

[4]Quotation marks around the word contempt indicate that we are referring specifically to the putative expression of contempt—the configuration of muscular movements that has been associated with the word *contempt* in previous research (e.g., Ekman & Friesen, 1986; Ekman & Heider, 1988).

dent events that can provoke it, even if they cannot produce a single emotion label to describe the expression.

This article presents two studies that address the methodological and theoretical questions raised in the above discussion. In the first study, we addressed the criticisms of Russell and others regarding the fixed-choice format by having subjects judge a set of facial expressions of emotion and provide their own labels for these stimuli. The first study is not only relevant to Russell's (1994) criticism about the whether the fixed-choice methodology makes use of the best labels for particular emotion expressions, but it also is relevant to the question whether contempt words are as salient as the words for other emotions. In Study 1 we also explored the extent to which contempt as well as other emotions are understood in terms of their relationships to antecedent events. In particular, we sought to determine whether agreement on "contempt" would be improved on the basis of relating it to an antecedent situation rather than an emotion label. Study 2 is a partial replication and extension of Study 1.

STUDY 1

Study 1 is presented in two parts. Part I deals with the methodological issue of whether the words or categories of words obtained from subjects' free responses are consistent with those presented in fixed-choice studies. Hypothesis 1 predicted that, for all facial expressions of emotion presented *except* contempt, when subjects provided their own label for each expression, they would provide the same label or one that is closely related to the label that has been associated with that expression on the basis of previous fixed-choice research.

Part II presents the data on two fixed-choice judgement conditions: story and label. In the story condition, subjects chose from a list of brief descriptions of events the one that was most likely to have produced the emotion expressed in each stimulus slide. In the label condition, subjects judged the slides using a typical, single label fixed-choice procedure. We hypothesized that, for the expression of contempt, agreement levels from both the story and label conditions would be significantly greater than agreement from the free-choice condition in Part I (Hypothesis 2), and that agreement levels from the story condition would be significantly greater than agreement levels from the label condition (Hypothesis 3). further, Hypothesis 4 posited that, for the other six emotions, agreement levels from the story condition would *not* be superior to agreement levels from the label condition. Table I lists all the hypotheses presented in this paper.

Table I. List of Hypotheses for Study 1 and Study 2

1. For all facial expressions of emotion presented *except* contempt, when subjects provided their own label for each expression, they would provide the same label or one that is closely related to the label that has been associated with the expression on the basis of previous fixed-choice research.

2. For the expression of contempt, agreement levels from both the story and label conditions would be significantly greater than agreement from the free-choice condition in Part I.

3. For contempt, agreement levels from the story condition would be significantly greater than agreement levels from the label condition.

4. For the other six emotions, agreement levels from the story condition would *not* be superior to agreement levels from the label condition.

5. Even when subjects had the opportunity to say that none of the listed antecedent situations fitted the facial expression, agreement levels for all emotions *including* contempt would be significantly greater than chance and comparable to those in previous studies.

Method

Subjects

Ninety-five undergraduates from San Jose State University and San Francisco State University participated in Study 1. There were 68 women and 27 men, between the ages of 18 and 63 ($M = 27.12$, $SD = 8.02$). All subjects were born in the United States and spoke English as a first language.

Design

We ran the study in two groups of approximately the same size—one at each campus. Within each group, subjects were randomly assigned to one of the three possible judgment conditions: free-choice, or one of two possible fixed-choice conditions (story or label). In the label condition, subjects chose from a list of seven emotion words; in the story condition they chose from a list seven antecedent event stories. The free-choice results are presented in Part I, and the label and story findings will be explained in Part II.

Procedure

The experimenter (E.L.R.) informed the members of each group that they would see a series of slides of facial expressions and that their task would be to judge what the person in each slide was feeling. Subjects completed a brief demographic questionnaire, after which they read the in-

structions at the top of their response form. After six practice trials, subjects began the experiment. Pilot studies indicated that subjects needed up to six practice trials to become accustomed to using the stories to judge slides in the story condition. In order to keep conditions constant across subjects and given the fact that all conditions were run together, those assigned to the free-choice and label conditions did the practice trials as well. Subjects received no feedback on their performance during the practice period.

Each slide was presented for 30 sec, during which time the subjects made their judgements. Each subject judged all 20 slides. After the last slide viewing period, the experimenter collected the forms and debriefed the group.

Stimuli

Subjects viewed a series of 20 slides: two for each emotion (anger, contempt, disgust, fear, happiness, sadness, and surprise) plus six practice slides (two slides of happiness, and one each of fear, surprise, contempt, and disgust). The slides were presented in a randomly determined order, which was the same for all subjects. Although it is certainly possible that slide presentation order has an effect on judgments, whatever effect it might contribute is minor relative to the effect we are studying here (i.e., agreement on the category of emotion judgment). A large corpus of studies on facial judgment have employed different (single) random orders and provided a consistent pattern of findings on the category of judgments (e.g., Ekman & Friesen, 1986; Ekman et al., 1987; Ekman, Sorensen, & Friesen, 1969). Additionally, our pilot studies employed two random orders of stimuli, and we found no differences on categorical agreement between order conditions.

Selection of Stimulus Slides. The slides of all emotions except contempt were taken from the *Japanese and Caucasian Facial Expressions of Emotion (JACFEE)* set of facial expression slides (Matsumoto & Ekman, 1988). From the JACFEE set, we used the four Caucasian examples of each emotion (except contempt).[5] The two best examples of each emotion were then selected on the basis of facial scoring (explained below). The contempt slides were not JACFEE slides. We shot our own contempt slides (for reasons explained below), and then selected the two best slides on the basis of facial coding.

[5]The JACFEE set includes slides of Caucasian and Japanese posers. We used only the Caucasian slides for this study because previous research has indicated that differences in the physiognomy of Japanese and Caucasian faces influences the perceived intensity of certain emotions (Matsumoto and Ekman 1989). As our new contempt photos were of Caucasian posers, we kept the ethnicity of the posers constant across all emotion expressions.

Facial Scoring for Slide Selection. We scored all potential stimulus slides using Ekman and Friesen's (1978) Facial Action Coding System (FACS). FACS dissects all observable facial movement into 44 visually distinguishable and anatomically separate units of muscular action. These units are called action units. FACS was designed for scoring moving facial behavior from videotape, and the coder notes changes in facial muscular action. The normal procedure for using FACS on still photographs or slides is to compare the slide of the poser making a facial expression to a slide of the poser's neutral face, and then to infer muscular movement on the basis of deviations from neutral. This type of scoring does not include static features of the poser's facial structure that may contribute to its appearance; it only scores deviations from a neutral pose. This method of FACS coding still photographs or slides can be referred to as the *baseline-corrected method*, and it has been used to code previous well-known sets of facial expression slides [e.g., Ekman & Friesen's (1976) *Pictures of Facial Affect* and the JACFEE set].

In the slide selection for this study, however, we adopted a coding approach that to our knowledge has not been used before on a formal basis.[6] We used FACS to score the expression slides without reference to a neutral face. The rationale for this approach is simple. Subjects do not have the opportunity to compare a stimulus face with a neutral face; thus the inferences they draw may be based both on the poser's physiognomy (which baseline-corrected scoring ignores) as well as any changes in appearance that result from muscular contraction. We call the FACS coding of a photograph of a poser's expression without comparison to a photograph of the poser's neutral face the *non-baseline-corrected method*. Non-baseline-corrected FACS coding scores every observable *apparent* muscular contraction, even those that may be static features of the poser's neutral face, in order to quantify in action units (AUs) all of the facial muscular cues that the subject actually sees.

Reliability. All faces were coded using non-baseline-corrected FACS by two experienced FACS coders (the first author and another FACS coder from our laboratory who was not involved in the present research). Intercoder agreement ratios for each slide were calculated between the pair of coders according to the following ratio: number of AU agreements divided by the total number of AU agreements and disagreements. The mean agreement ratio was .95, $s = .07$, mode $= 1.00$. The few differences were arbitrated among three coders: the two original coders and the second author (an expert in facial coding).

---

[6]FACS scoring principles that considered static features of the poser were informally applied in previous research (Ekman et al., 1987), but the first systematic coding of this kind is reported here.

On the basis of the final, arbitrated codes, we selected the two best slides according to the following criteria: (1) We retained only those slides that contained the minimal AUs necessary for categorization as an example of a prototypical expression of each of the emotions in the study. To determine the degree of an expression's prototypicality, we ran the FACS scores (i.e., the AU-based descriptions of each expression) through a computerized program that determines whether AU combinations include core facial movements that characterize certain facial expressions of emotion. The program's interpretations draw on a rich empirically and theoretically derived database of expression data from our laboratory and others', and it has been used for the classification of spontaneous facial behavior in previous studies (e.g., Chesney, Ekman, Friesen, Black, & Hecker, 1990; Ekman, Davidson, & Friesen, 1990; Rosenberg & Ekman, 1994). (2) Expressions could not show evidence of any AUS that are relevant to other emotions. (3) If the slide passed the first two criteria, the next most important feature was that the stimuli matched on intensity. Each AU was scored for the intensity or strength of muscular contraction on the FACS 5-point intensity scale (Friesen & Ekman, 1992). We matched intensity within emotion (for any pair of slides of same emotion that passed criteria 1 and 2) and across emotions. With respect to the former, all AUs critical to the interpretation of an emotion had to agree within 1 intensity point on the FACS 5-point intensity scale. Also, overall intensities of expressions could not vary more than 1 point across all seven categories of emotion stimuli.

The two slides for each of the seven emotion categories that met all of the above criteria were selected to be included in final stimulus set. If more than two slides met the above criteria, then the best male and female poser were chosen. However, if the best two slides for a given emotion selected according to the above criteria were posers of the same sex, then they were the stimuli used. Previous studies on judgement of facial expressions have found no differences in judgments of emotion on the basis of poser sex (Ekman, 1989; Russell & Fehr, 1987).

Contempt Slides. Ekman and Friesen (1986) found that unilateral lip tightening was consistently recognized as contempt among more judges across cultures than two other proposed contempt expressions (e.g., bilateral lip tightening; or raising the entire upper lip slightly, without tightening of the lip corners). This configuration has been recognized as contempt in a replication study (Ekman & Heider, 1988) and in other cross-cultural research (e.g., Matsumoto, 1992). Some theorists and investigators of facial expression have suggested, however, that the expression of contempt contains a head tilt or turn and/or eyes averted or turned (Darwin, 1872/1965; Izard, 1971; Izard & Haynes, 1988). A study of decoding and encoding of

contempt in Northern and Southern Italians and Americans (Ricci-Bitti et al., 1989) revealed that extrafacial cues such as head and eye position may contribute to the identification of the unilateral lip tightening as contempt rather than disgust or happiness (emotions with which contempt is sometimes confused).[7] In particular, the head turned to the side or tilted up, along with the eyes looking to the side, or looking downward from an upward tilted head position may help convey the condescension of contempt.

The contempt slides in the JACFEE set include the facial actions mentioned above (unilateral lip tightening), but the head and eyes are center. Given the Ricci-Bitti et al. (1989) findings, we chose to shoot new contempt slides that combined the Ekman and Friesen (1986) configuration of unilateral lip tightening with minor head and eye position deviations: a slight head turn and eyes looking to the side or at the camera from the side. We shot several examples with two female and two male posers. The best four slides were then shown to a group of 51 psychology undergraduates, interspersed with other emotion slides (including contempt) from the JACFEE set. Subjects chose from a list of labels (anger, contempt, disgust, fear, happiness, sadness, and surprise) the word that they thought best described what the person in each slide was feeling. All of the new contempt slides showed higher judge agreement on contempt than the published norms on the JACFEE slides of contempt. We then selected the best examples of the new contempt slides on the basis of coding criteria listed above for inclusion in the experiment. The two best poses were by one female and one male poser, both of whom showed the highest agreement levels on contempt in the pilot study (67% and 71%, respectively).

Response Forms

The subjects in the free-choice condition (Part I) received a response form on which the instructions were: "Your task in this experiment is to look at the facial expression shown in each slide and make a judgment about how the person in the slide feels. In each blank space listed below, write the one word that you think best describes how the person feels." For Part II (comparison of the two choice conditions) subjects judged each slide using either the label response form or the story response form. We created two random orders for the emotion words or stories to appear on the forms.

---

[7]Pio Ricci-Bitti of the University of Bologna, our collaborator on a cross-cultural extension of this research, encouraged us to include head and eye variations in the contempt expression slides.

The label response form simply contained a brief description of the task (instructions printed below) and a list of seven emotion labels: anger, contempt, disgust, fear, happiness, sadness, and surprise. The instructions were:

> Your task in this experiment is to look at the facial expression shown in each slide and make a judgment about how the person in the slide feels. Listed below are several emotion words. Pick the one word that you think best describes how the person feels. Write the specific word that you think best describes what you think the person in each slide is feeling in the blank next to each slide number at the bottom of this page. Please read through the list of words before we begin.

The instructions were modified slightly for the story response form:

> Your task in this experiment is to look at the facial expression shown in each slide and make a judgement about how the person in the slide feels. Listed below are several stories about events that can lead to certain emotions. Pick the story below that would be most likely to produce the expression shown in each slide, if the person on the slide were "the person" described in the story. To the left of each story is a name for the story. Write the name of the story that would be most likely to produce the emotion shown in each slide in the blank next to each slide number at the bottom of this page. Please read through each of the stories before we begin.

Development of the Emotion Stories

For each of the seven emotions that were listed as labels and shown in the slides, we developed brief (one- or two-sentence) scenarios that described an emotional event that happened to a fictitious person. This approach to obtaining emotion judgments, developed by Dashiell (1927), was used in Ekman and Friesen's (1971) early work in New Guinea and in studies of children's understanding of facial emotion (Holder & Kirkpatrick, 1991). In developing our stories, we referred to dictionary definitions of the emotion words and various emotion theorists' definitions of these words. The central themes for each story are most similar to those described as common emotion-eliciting situations by Ekman (Ekman, 1977; Ekman & Friesen, 1975), except for the happy story, which is similar to the antecedent conditions described in Scherer and Wallbott's (1986) cross-cultural study of reported emotional antecedents. The themes described by Ekman (1977) for anger, fear, and sadness are all consistent with the empirical data on these emotions reported by Scherer and Wallbott, which lends credibility to their use here. Rozin (personal communication, 1992) has used a contempt story that is similar to ours with success in a task in which faces were matched to antecedent events. None of the stories contained affective-laden terms or gender-referential pronouns. We developed

several stories for each emotion and then selected the best exemplars for each emotion on the basis of extensive discussion.

We administered the stories to a group of 40 psychology undergraduates at San Jose State University. Their task was to choose from a list of words (anger, contempt, disgust, fear, happiness, sadness, and surprise) the word that they thought best described the emotion that the person in each story might have felt. Subjects were allowed to use a word more than once if they felt it was necessary. For all of the stories, the percent agreement with the predicted emotion was significantly greater than chance (based on binomial tests that used the same values for chance as the first criteria reported in study 1, $p < .01$ for all stories). The seven stories are listed in the Appendix. On the story response forms in the experiment described below, subjects simply wrote in the name of each story (listed to the left of each story in the Appendix) that best described what might have just happened to the person in each slide. They were not given the emotion labels that fit each story.

Results and Discussion

Part I: Agreement in the Free-Choice Condition (n = 35)

Categorization of Free Responses. The words from the free-choice judgments were categorized into emotion and non-emotion categories using a lexicon of emotion words developed in the Human Interaction Laboratory (Ekman & Irwin, 1994). The lexicon is a database of 508 root words classified into 28 categories. The lexicon's nonredundant category list was derived from all emotion categories reported by past investigators in the study of facial expression of emotion judgements, categories proposed by discrete emotion theorists, and categories that were developed in the process of classification of recent judgment data (not from this experiment). The words in the lexicon were derived from a large, unpublished data set of 200 college students' free responses to 64 photographs from Silvan Tomkins' series of facial expressions. This data set was collected by Robert Krause and Seymour Rosenberg in the 1960s. Ekman and Irwin assigned words from the Krause and S. Rosenberg data set into the lexicon's categories on the basis of the definition of each word in the *American Heritage Dictionary (Second College Edition)* (1985).

We classified each word from the present data set into the category to which the same word belonged in the lexicon. For words not included in the lexicon, the two authors made independent categorizations (blind to knowledge of the slides for which they occurred) and then arbitrated any

Table II. The Words Given by Subjects in Each Predicted Category of Facial Expression (Categories Listed Are for Each Stimulus Expression, Which Correspond to the Lexicon's Word Categories)[a]

Anger	%	Contempt	%	Disgust	%	Fear	%
Angry	25.8	Sarcastic	7.4	Disgust	34.3	Fear	28.6
Anger	15.7	Smug	5.9	Disgusted	17.1	Scared	27.1
Mad	18.6	Cynical	2.9	Dislike	4.3	Frightened	4.3
Frustrated	8.6	Sarcasm	1.5	Grossed-out	2.9	Afraid	2.9
Frustration	5.7			Disgust (upset)	1.4	Fright	2.9
Furious	4.3			Smells bad	1.4	Anxiety	1.4
Fury	2.9					Fearful	1.4
Pissed	1.4					Terror	1.4
Rage	1.4						
Fighting	1.4						

Happiness	%	Sadness	%	Surprise	%
Happy	77.1	Sad	42.9	Surprise	35.7
Happiness	11.4	Sadness	12.9	Surprised	24.3
Pleased	2.9	Depressed	4.3	Amazed	4.3
Friendly	1.4	Blue	1.4	Astonished	2.9
Good	1.4	Forlorn	1.4	Amazement	1.4
		Pity	1.4	Awe	1.4
		Sad and thoughtful	1.4	In Awe	1.4
		Saddened	1.4	Oops-surprise	1.4
		Self-pity	1.4		
		Unhappy (sad)	1.4		

[a]Numbers listed are percentages of total number of responses, $n = 70$ (2×35 subjects).

disagreements. We consulted the American Heritage Dictionary (Second College Edition) (1985) to resolve any final disagreements.

Were Expressions Consistently Recognized as the Predicted Emotions When Judges Provided Their Own Labels? Table II presents all of the words given within the predicted category for each type of facial expression. The predicted category was the modal response category for all emotions except contempt, the various responses for which will be discussed in more detail below. For all emotions except contempt, most subjects gave the exact word (or a close variant) as those that have been used in most fixed-choice studies (c.f. Ekman et al., 1987).

We combined the percentages of all words within the predicted categories from Table II, as classified by the lexicon, into a collapsed category percentage that is shown in Table III. As mentioned earlier, there were two slide stimuli for each emotion. The percentages for each exemplar (poser) for each emotion did not differ for any of the emotions. The percentages presented in Table II are the percentages of the total number of responses given to each category of emotion slide, summed across the two

Table III. Percentage Agreement of Free-Choice
Categories with Predicted Emotions for Seven
Facial Expressions in Study 1, Part I ($n = 35$)

Emotions		Chance Criterion 1	Chance Criterion 2
Anger	85.7[a,b]	.20	.33
Contempt	17.7	.20	.33
Disgust	61.4[a,c]	.20	.33
Fear	70.0[a]	.20	.50
Happiness	94.3[a]	.50	N/A
Sadness	70.0[a]	.20	N/A
Surprise	71.0[a,c]	.33	.50

[a]Different from Chance Criterion 1, $p < .001$.
[b]Different from Chance Criterion 2, $p < .001$.
[c]Different from Chance Criterion 2, $p < .01$.

slides, for a total of 70 responses (there were 35 subjects, and each subject judged two slides for each emotion). For the statistical tests presented in Table III, however, we collapsed across both slides for each emotion by taking the *mean agreement level* across the two slides. This approach to collapsing judgments preserves the original number of subjects in calculation of the test statistics, thereby not violating assumptions of independence (c.f. Wagner, 1993).

We then tested whether each proportion listed in Table III differed from what would have been expected by chance using binomial tests. Two types of chance agreement levels were derived. The first chance agreement level (Chance Criterion 1) was defined as the probability of choosing "x" emotion out of all slightly similar emotions. For all of the clearly negative emotions (anger, contempt, disgust, fear, and sadness), Chance Criterion 1 was the probability of any one of them being chosen: 1/5 or .20. For happiness, the only positive emotion on the list, the most conservative chance criterion reflected the choice of a positive versus a negative expression, which is .50. In the case of surprise, we used a chance criterion of 1/3 to reflect the fact that while most people report surprise for this expression, some report either a positive expression (happiness) or a negative expression (fear) with which this ambiguously valenced emotion is occasionally confused (Tomkins & McCarter, 1964).

For all emotions except happiness and sadness, we tested the observed agreement levels against a second, more stringent chance criterion. Chance Criterion 2 reflected the probably of picking "x" emotion out of a set of expressions with which it is commonly confused or shares morphological characteristics. Anger, contempt, and disgust are expressions that share

some morphological as well as semantic features. Thus, for each of these facial expressions, Chance Criterion 2 was the probability of choosing any one of them randomly: 1/3 or .33. Also, fear and surprise share morphological characteristics with each other (raised eyebrows, wide eyes), so Chance Criterion 2 for both of these expressions was 1/2 or .50. We did not test sadness against a second criterion, because there are no clear options for another facial expression with which it is confused. The chance criterion for happiness could not be made more stringent; therefore we did not derive a second criterion for this emotion.

All of the facial expressions except contempt were recognized as the predicted emotions (i.e., the emotion category labels that have been used for these expressions in previous fixed-choice studies) significantly more often than would have been expected by the first chance criterion. The findings support Hypothesis 1, which predicted nonchance agreement for all emotions except contempt. All of the noncontempt emotions except fear were recognized at levels greater than the more stringent Chance Criterion 2. Although fear agreement was substantially larger than the very stringent second chance criterion, and the p-value for the comparison of the observed level of agreement with Chance 2 was .13.

Some might disagree with the inclusion of frustration words in the anger category (e.g., Russell, 1994).[8] We recalculated the percentage agreement for the anger category excluding the frustration words and the agreement is 71.5%, which is greater than both chance criteria at the .01 level.

The results for the disgust expression, though significantly better than chance, were lower than those for the other emotions. Analyses of the alternate responses for the disgust expression show that 29% of the judges gave a word in the lexicon's anger category. The exact words were: *angry, anger, mad, bitter, snarly, hate, frustrated, frustration,* and *rage.* Anger is one of the emotions with which disgust is commonly confused (Ekman, 1972; Tomkins & McCarter, 1964), which might explain this response pattern. Subjects did not give contempt words for the disgust expression, even though contempt is also a "common confusion" for disgust. This might be a result of the infrequent usage or poor understanding of the contempt label; after all, very few subjects provided a word in the contempt category in response to the "contempt" slides.

[8]One of the anonymous reviewers questioned the inclusion of *frustrated* and *frustration* in the anger category, reminding us that Russell (1994; Russell & Fehr, 1994) has argued that frustration is not an emotion. We included these words in the anger category because a good deal of research and theory on emotion show that frustration is an antecedent of anger (c.f. Arnold & Gasson, 1968; Ellsworth & Smith, 1988; Lazarus, 1991; Weiner, 1985). Russell's own research on anger words (Russell & Fehr, 1994) also indicates that people categorize frustration as a type of anger, even if it is not a prototypical form.

Was There Agreement on the Freely Chosen Labels for the Contempt Expression? Contempt agreement was not significantly greater than either chance criterion. This is consistent with our first hypothesis, which predicted that agreement on the predicted label for the "contempt" expression would not be greater than chance. How did people label the "contempt" expression? Only 17.7% of the subjects provided a word that the lexicon classifies in the contempt category. The exact responses in the contempt category are listed in Table II. The modal response category (26%) was a wide variety of words in the lexicon's category of Attention-Thought: *unsure, doubtful, skeptical, uncertain, concentrate, contemplative, pensive, bored,* and *doubt.* Twenty percent of the judges gave the following anger category words: *annoyed, pissed-off, evil, anger,* and *irked.* Sixteen percent of the responses were for the Emotion, Unspecific category: *disbelief* and *incredulous.* None of the agreement levels from any of these response categories was significantly greater than either chance criterion.

What are we to make of the contempt findings? We have made the argument that the low agreement on the "contempt" expression may occur because the word *contempt* is either infrequently used, poorly understood, or underrepresented in the common vernacular. An alternative explanation is that while the concept of contempt is well understood, people simply do not recognize the "contempt" expression to convey the emotion of contempt. Such an account, however, is not consistent with fixed-choice results from within- and cross-cultural research on this expression (e.g., Ekman & Friesen, 1986; Ekman & Heider, 1988; Matsumoto, 1992). In that work, people consistently have chosen the label *contempt* for the expression of "contempt," which suggests that they have an understanding of the emotion. The poor free-choice data might simply indicate that people either do not know or cannot retrieve the word *contempt,* even if they understand the social implications of the emotion. We undertook Part II of Study 1 to specifically determine whether people do understand the "contempt" expression in terms of an antecedent event that might elicit it.

Part II: Comparison of Agreement Between the Two Fixed-Choice Conditions (n = 60)

For the analyses presented in Part II, we collapsed across the two stimulus slides for each emotion using the same procedure described in Part I, because there were no systematic differences between posers. Also, there were no differences between the two emotion orders for each response form, so we collapsed across these different orders within each type of response form.

Table IV. Percentages Agreement on Predicted Emotions for Seven Facial Expressions in Two Fixed-Choice Conditions, Study 1, Part II

Emotions	Label ($n = 31$)	Story ($n = 29$)	z^a
Anger	96.8	96.6	0
Contempt	75.4	94.6	-2.18^c
Disgust	75.8	87.9	-1.23
Fear	83.9	91.3	-0.83
Happiness	96.7	96.6	0
Sadness	93.6	100	-0.64
Surprise	91.9	67.2	2.50^b

[a] All tests two-tailed except for contempt, for which we made a directional prediction.
[b] $p < .05$.
[c] $p = .01$.

Did Agreement for Contempt Improve over Free-Choice Levels? Binomial tests showed that agreement levels for all emotions for both conditions were significantly greater than chance at at least the .01 level, using the same chance criteria described in Part I and listed in Table III (agreement levels for the story condition for surprise was only greater than Chance Criterion 1 at the .01 level, it was not significantly greater than Chance Criterion 2). The fact that the agreement levels for contempt in the story condition (as well as the label condition) were significantly greater than both chance criteria strongly supports Hypothesis 2, because the percentage agreement on contempt in the free-choice task from Part I (17.7%) was less than either of the chance criteria.

Was There a Difference in Percentage Agreement Across the Two Response Conditions? Hypothesis 3 posited that, for the expression of contempt, agreement levels for the story condition would be greater than those in the label condition, while Hypothesis 4 predicted that agreement levels would not differ between conditions for the other emotions. To test these hypotheses, we compared the percentage agreement between subjects on the predicted emotional response across the two response conditions (e.g., whether the subjects chose the anger label or the anger story for the slide which we picked a priori as an expression of anger). We conducted z-tests of the differences between independent population proportions (Guilford, 1954; Shott, 1990) separately for each emotion. The results are presented in Table IV.

For contempt, agreement improved significantly in the story condition, as predicted by our third hypothesis. For most other emotions, percentage agreement with the predicted interpretation did not differ between conditions, which supports Hypothesis 4. Subjects performed at least as well in the story condition, except for the expression of surprise.

We did not predict that the agreement on surprise would be significantly lower in the story condition. An agreement level of 67.2% is, however, greater than what would have been expected by chance ($p < .01$) when chance is .33 (based on the notion that the emotion surprise is most likely to be confused with is fear, and sometimes happiness, so the chance of picking any one randomly is 1/3). Analysis of alternative responses showed that 19% of the subjects chose the fear story for the surprise face. The high levels of agreement in the label condition for both surprise and fear—and the fact that subjects did not pick the surprise story for the fear face—indicate that something may be wrong with the surprise story in conveying the antecedent conditions of this emotion. With the exception of surprise, however, the results indicate that, when people refer to emotional situations to make their judgments of facial expressions, they do at least as well as judges who simply choose verbal labels. Certainly all of the agreement levels across both conditions are comparable to what has been previously published.

The high levels of agreement on the "contempt" expression in the story condition indicate that people do recognize emotional meaning in the facial expression, and that the situational precursors of this expression might be more accessible than its verbal label. But was the statistically significant difference between the story and label conditions for contempt a substantial one? This can be addressed by looking at the effect size of the difference between the story and label agreement proportions for contempt. We calculated the difference between the arcsine transformations of these proportions to calculate Cohen's h (Rosenthal & Rosnow, 1991, p. 449). The difference between conditions indicated a moderate to large effect ($h = .68$), which underscores the importance of this finding. The fact that agreement levels of most of the emotions from both of the fixed-choice conditions in Part II were greater than those obtained from free-choice in Part I will be discussed in the general discussion.

STUDY 2

One could question whether the superior agreement levels for contempt in the story condition were an artifact of exclusion. That is, subjects may have chosen the contempt story simply because none of the other antecedent situations fit the "contempt" expression. This has been one of Russell's (1994) and Wierzbicka's (1986) criticisms of the fixed-choice methodology with single labels.

In order to determine if the antecedent event agreement levels for contempt and other facial expressions of emotion might be a result of exclusion as well as to replicate the effects of the antecedent event-based judgments,

we repeated the story condition of Study 1 with one important modification. This time we gave observers the opportunity to say that none of the stories adequately described what the person in the slide was feeling; i.e., it was no longer a *fixed*-choice condition. If we obtained nonchance agreement levels in this paradigm, then we could be confident that our fixed-choice story condition results for all emotions from Study 1 were not due to exclusion. This issue was most urgent for the expression of "contempt," which fared so poorly in the free-choice task of Study 1. Ruling out exclusion as a basis for the high agreement in the story condition would strengthen our interpretation that contempt is well understood in terms of the situations that call it forth.

We predicted that, even when subjects had the opportunity to say that none of the listed antecedent situations fitted the facial expression, agreement levels for all emotions *including* contempt would be significantly greater than chance and comparable to those in previous studies (Hypothesis 5). We did not rule out the possibility, however, that agreement levels might be somewhat attenuated from the levels is shown in Study 1, for observers would not be forced to commit to any emotion story in their judgements.

Method

Subjects

Thirty-one undergraduates from San Francisco State University participated in Study 2. There were 21 women and 10 men in the sample, ranging in age from 18 to 45 ($M = 24.77$, $SD = 5.30$). All subjects were born in the United States and spoke English as a first language.

Procedures and Design

We ran the study in one large group in a classroom at San Francisco State University. There were two design modifications: We ran only the story condition, and this time subjects had the opportunity to choose none of the stories. This was explained in the instructions on the response forms. Otherwise, we conducted the experiment just as the Study 1. Once again, each subject judged all 20 slides.

Stimuli

We used the same set of slides from Study 1 and presented them in the same order.

Table V. Percentage Agreement on Predicted
Emotions for Seven Facial Expressions Using the
Story Response Form in Study 2 and Comparison
with Study 1, Part II, Levels ($N = 31$ for Study 3)

Emotion	Percentage agreement	z for comparison with Study 1, Part II
Anger	80.3[a,b]	-2.06[e]
Contempt	93.5[a,b]	-0.18
Disgust	72.5[a,b]	-1.53[d]
Fear	82.0[a,c]	-1.07
Happiness	85.2[a,b]	-1.58[d]
Sadness	88.5[a,b]	-2.01[e]
Surprise	45.0	-1.78[e]

[a]Significantly different from Chance Criterion 1, $p < .001$.
[b]Significantly different from Chance Criterion 2, $p < .001$.
[c]Significantly different from Chance Criterion 2, $p < .01$.
[d]$p < .10$, one-tailed test.
[e]$p < .05$, one-tailed test.

Response Forms

The forms were similar to the story response forms used in Part II of
Study 1, except we modified the instructions slightly and created a new
response option. The instructions read just as in Study 1, except an addi-
tional sentence was added that said: "If there is no story that fits what the
person in the slide is feeling, then write 'no fit' in the blank." The label
no fit appeared in the list of story names, next to a brief sentence: "No
story fits what the person in the slide is feeling." Once again we created
two forms, in which there were two possible orders of the emotion stories.

Results and Discussion

Were the Agreement Levels from Study 1, Part II Maintained?

As is the previous two studies, there were no differences in percentage
agreement between the two slides for each emotion, so we collapsed the
scores across slides. Nor were there differences among the two forms (or-
ders) for the stories, so we collapsed across the alternative response forms.
The first column of Table V presents the agreement levels on each of the
predicted emotion categories. We conducted binomial tests to determine
if the agreement levels for each emotion differed from those that would

have been expected by chance and used the same chance criteria for each emotion as those used in Study 1, Part I. For all emotions except surprise (to which we have already attributed problems with the story), agreement levels were still far greater than chance levels would have predicted, even when they were evaluated against the more stringent chance criteria.

Did Judges Still Agree on Contempt When Exclusion Could Be Ruled Out as a Basis for the Choice?

Even when observers had the opportunity to say that none of the stories adequately described what the person in the "contempt" slide was feeling, 93.5% chose the contempt story. This result offers strong support for Hypothesis 5. Consensus on the story that elicits the "contempt" expression cannot be attributed to subjects choosing that story simply because none of the other options applied. In fact, story-based agreement levels on contempt did not change significantly from the story condition in Study 1 (as explained below), which strongly suggests an association between the "contempt" expression and the antecedent situation represented in the contempt story.

Comparisons with Study 1: What Were the Effects of Exclusion on Agreement?

As a *post hoc* statistical comparison between the agreement levels in this study (to which exclusion principles presumably do not contribute) and those from Study 1, Part II, we conducted z-tests of the differences between proportions for each study. The z-statistics for these comparisons are listed in the second column of Table V. For anger, sadness, and surprise the agreement in Study 2 was significantly less than in the story condition of Study 1. The decreases for fear, happiness, and disgust were not significant, nor was the trivial decrease for contempt.

It is not surprising that overall levels of agreement decreased slightly in some instances. The option of saying "no fit" may have capitalized on any ambiguity subjects may have had about the emotional meanings of the stories. Still, the levels of agreement observed in this study were consistent with levels obtained in the corpus of research on facial emotion judgment— all were in the 70s to 80s (the exception of surprise is discussed below).

Considering that the agreement levels for the story-based judgments of several of the other emotions decreased slightly from Study 1, Part II, it is even more remarkable that the agreement for the story-based judgments of "contempt" did not decrease on a substantive or statistical basis.

This finding strengthens the interpretation that people understand the situational concomitants of the "contempt" expression, even if they cannot freely provide the predicted label for it.

The Problem of the Surprise Story

Once again we obtained very low agreement for judgments of the surprise expression using the antecedent story. This time the alternative responses were split between fear (20%) and no fit (18%). Fear and surprise expressions are sometimes confused, though results in previous studies have indicated that the amount of confusion should not be this high. As we mentioned in the results to Study 1, it is quite possible that the surprise story is weak, and thus subjects would have chosen the next most viable option for the expression. Fear is a logical choice, as it shares many morphological characteristics with surprise, including raised eyebrows, wide eyes, and an open mouth (Tomkins & McCarter, 1964). The fact that 18% chose no fit for the surprise expression in the story condition, however, further indicates that the surprise story might not adequately describe a surprise elicitation situation. Ekman (1977) stated that "surprise elicitors share the characteristics of being unexpected, novel, and are usually sudden rather than gradual" (p. 60). Our story may not have captured the suddenness necessary for surprise, but we can only speculate on this at this time. Given the high agreement on surprise shown in the free-choice task of Study 1, Part I (74.3%), as well as the high agreement on surprise in the fixed-choice label condition for Study 1, Part II (91.9%) and in numerous other studies (c.f., Ekman et al., 1987), it is likely that the poor agreement for surprise in the story conditions across Study 1 and Study 2 is more suggestive of an inadequate story than it is of a poorly understood facial expression.

GENERAL DISCUSSION

The present research addressed two key questions about the judgment of facial expressions of emotion; one methodological, the other conceptual. We sought to determine whether a free-choice methodology would provide evidence of agreement on the same labels or similar labels to those used in previous fixed-choice studies. A further interest was whether subjects could use situational antecedents of particular emotions to identify facial expressions of emotions; and in particular, whether this approach would improve agreement levels for the lexically ambiguous emotion of contempt. Our findings provide affirmative answers to both of these questions.

In the following sections, we discuss the important methodological and conceptual implications for research on the understanding of facial expressions of emotion raised by the present findings. The methodological issues focus on the validity of word lists used in most fixed-choice research as well as the relative stability of facial judgments across a variety of response formats. The conceptual issues derive from the findings on contempt, and concern the relationship between the lexical and situational representation of emotion as conveyed by facial expressions.

Methodological Implication 1: Fixed-Choice Categories Are Supported by Free-Choice Data

For the judgment of all expressions except that of contempt, we obtained evidence that people not only freely provide labels in the same category of emotion as those categories used in fixed-choice studies, but a majority of them also provided the exact label that has been used in most fixed-choice studies. Russell (1994) argued that, if emotion recognition were categorical and specific, "then a list of emotions supplied by a researcher should not be necessary for recognition to occur" (p. 118). We have shown that a list is not necessary; recognition of certain facial expressions of emotion is categorical and specific. Lists are methodological conveniences, ones that we will continue to use along with other methodologies, because our findings demonstrate that they accurately represent the verbal categories people use when they freely label facial expressions of emotion in judgment tasks.

Methodological Implication 2: There Is Evidence For Agreement on Facial Expression Emotion Judgments Across Three Response Formats

We employed various techniques for obtaining facial judgments in the present research, which can also be evaluated as replications of the consistency in judgments of the same stimulus expression using at least three response formats: free-choice, fixed-choice using labels, and fixed-choice using antecedent stories. The free-choice condition—probably the most stringent response format as it imposed the least structure—yielded agreement levels consistent with those previously reported for all emotions except contempt. In all of the other conditions, although there were minor variations, all agreement levels were significantly greater than those that would have been expected by chance, and not grossly different from each other. The demonstration of high, nonchance agreement levels across the

studies strongly contradicts Russell's claim that "as we move from more restrictive to less restrictive attempts to capture the observer's response, a different interpretation of that response emerges" (1994, p. 123).

We agree with Russell that it is important to specify the extent to which methodology can influence judgments of facial emotion. Where we diverge is in the significance we attribute to the changes in agreement levels across methods. We have demonstrated that, with the exception of contempt, variations across methodologies are minor, and that consensus levels are comparable with those shown in the large number of fixed-choice studies. Russell (1994) has found greater variation across methodologies than we did, but we attribute this to a few key methodological differences between our approach and his. There are at least three reasons why Russell's growing body of studies on facial judgment do not adequately assess method variance.

First, he has addressed this issue through a variety of studies, many of which use different sets of stimuli. It is crucial to use the same set of stimuli across studies when one is studying the effects of response format only, otherwise variations in agreement levels due to minor differences in stimulus slides cannot be separated from variations due to judgment methodology. We have not only held our stimuli constant, but we selected our slides via rigorous procedures that ensured consistency in configuration within emotion and consistency in intensity within and across emotions.

Second, we have exercised greater control in obtaining judgments. In most of the studies that Russell has presented as evidence of low agreement in emotion judgments, his experimenters personally approached people in public places, showed them pictures for as long as they needed to see them, and then asked them to rate the pictures (Russell, 1991a, 1991b, 1991c; Russell & Fehr, 1987). While we appreciate Russell's interest in making the judgmental context less confined than ours (in which slides were shown for prescribed periods of time, and judges' responses were made privately without direct personal contact with the experimenter), his methods are likely to contribute to the method variance that he was intent on isolating.

Third, recently Russell (1991a, 1991b, 1991c) has conducted a series of studies on contempt recognition, and has presented inconsistency in agreement on *this* emotion as adequate evidence that there is no universal recognition of facial expression of *any* emotion. We have provided data indicating that while contempt expressions are consistently recognizable across at least two methodological techniques—and cannot be attributable to exclusion artifacts of the fixed-choice methodology—there are some problems with the labeling of this emotion that do not plague anger, disgust, fear, happiness, sadness, and surprise. Thus, contempt is not a fair test case of the extent to which facial expressions of most of the basic emotions are consensually understood.

*Conceptual Implications: There is Greater Consensus on the Si-
tuational Antecedents of the Contempt Expression than on the
Verbal Label for the Contempt Expression*

We found that people can consistently identify the "contempt" expres-
sion on the basis of an emotional antecedent, that agreement using this
procedure is significantly greater than the already respectable single-word
fixed-choice levels, and that antecedent event-based agreement cannot be
attributable to a strategy of ruling out by exclusion. The fact that agreement
levels for the contempt expression were significantly higher in the condition
in which subjects used emotional scenarios rather than labels for identifi-
cation contradicts the position that " . . .Ekman & Friesen's (1986) result
is limited to that subsample of human beings who know the word *con-
tempt* . . ." (Russell, 1991a, p. 163).

Ekman (1994) has proposed that "[T]he matching of words and facial
expressions is imperfect, at least in part, because they each convey what the
other cannot" (p. 270). The free-choice findings on the "contempt" expression
certainly indicate that the word-expression linkage is looser for this emotion
than it is for the other six that we studied. We have argued that *contempt* is
either poorly represented in the common vernacular of American college stu-
dents (albeit of varied ages), infrequently used, or both. Presently we cannot
distinguish among these alternative explanations, but the inaccessibility of this
word undoubtedly contributes to its weak link with the expression.

In general, the difference between agreement levels between fixed-
choice and free-choice approaches might simply be a function of the dif-
ferences between recognition and recall memory. According to Anderson
(1985), recognition is easier than recall memory at least in part because
recognition memory offers more access routes into a memory network. In
free-choice (a recall) task, the subject is only able to activate his or her
own network of knowledge of that emotion via a facial route, whereas in
a fixed-choice task (which is most likely a recognition memory task), sub-
jects have a verbal or situational as well as facial entrée into the network.
In the case of contempt, the network might not be as well established as
it is for the other six expressions. Subjects may have good access to the
facial and antecedent nodes in a memory network for that emotion (as the
story results would imply), but the word *contempt* might not be well enough
represented in the verbal node of the network to be accessed in a recall
task. The fixed-choice findings for "contempt" could be explained by the
word *contempt* priming semantic aspects of the network enough to facilitate
recognition of the face.

Presently, we are investigating whether agreement levels on the ante-
cedent story for recognition of the "contempt" expression are better than

those from the single-label fixed-choice methodology in Italian subjects, to determine if the lexical problem with contempt is unique to American English, or if it is a more general problem.[9] In our culture, at least, the facial expression of contempt has an unusual status—people know what provokes it, but most do not know what to call it. As far as we know, this cannot be said of any other facial expression of emotion that has demonstrated reliable cross-cultural recognition value. If we can determine why the expression of contempt is more readily labeled with a situation than a word, then perhaps we can understand the extent to which recognition of facial affect is related to more general knowledge about emotions and their provocations.

APPENDIX

Emotion Stories[a]

Story Name	
BRAKES	The person has realized that the brakes don't work while driving down a steep hill. The car approaches the end of the road, which is a cliff with no barrier. The person tries to brake and veers out of control.
FRIENDS	The person sees many close friends at a party.
DOG	The person steps in dog feces, reaches down to wipe it off, and feces gets on the person's hand.
POST OFFICE	The person is waiting in line at the post office for a very long time. The person finally reaches the window, when the clerk announces that there is time for only one more customer. The person is then pushed aside when someone cuts in front to take the person's turn.
TALL	The person is sitting next to someone who suddenly stands up and is much taller than the person had expected.
TAKE CREDIT	The person hears as acquaintance bragging about accomplishing something for which the acquaintance was not responsible.
CHILD	The person remembers the time last year when a young child died of a terminal illness.

[a]Observers were instructed to write in the name of the story that best described what might have just happened to the person in each slide. They were not given the emotion labels that went with each story. (Key: BRAKES = fear, FRIENDS = happiness, DOG = disgust, POST OFFICE = anger, TALL = surprise, TAKE CREDIT = contempt, CHILD = sadness).

REFERENCES

American Heritage Dictionary (Second College Edition) (1985). Boston: Houghton-Mifflin.
Anderson, J. R. (1985). *Cognitive psychology and its implications* (2nd ed.). New York: W. H. Freeman.

[9]With Pio Ricci-Bitti of the University of Bologna.

Arnold, M., & Gasson, J. A. (1968). Feelings and emotions as dynamic factors in personality integration. In M. Arnold (Ed.). *The nature of emotion* (pp. 203-221). Harmondsworth, England: Penguin.

Chesney, M. A., Ekman, P., Friesen, W. V., Black, G. W., & Hecker, M. H. L. (1990). Type A behavior pattern: Facial behavior and speech components. *Psychosomatic Medicine, 53*, 307-319.

Darwin, C. (1965). *The expression of the emotions in man and animals.* Chicago: University of Chicago Press. (Original work published 1872).

Dashiell, J. F. (1927). A new method for measuring reactions to facial expressions. *Psychological Bulletin, 24*, 174-175.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation, 1971* (pp. 207-283). Lincoln: University of Nebraska Press.

Ekman, P. (1977). Biological and cultural contributions to body and facial movement. In J. Blacking (Ed.), *Anthropology of the body* (pp. 53-84). London: Academic Press.

Ekman, P. (1989). The argument for and evidence about universals in facial expressions of emotion. In H. Wagner & A. Manstead (Eds.), *Handbook of social psychophysiology* (pp. 143-164). London: John Wiley & Sons.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115*, 268-287.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology II. *Journal of Personality and Social Psychology, 58*, 342-353.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*, 124-129.

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

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

Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motivation and Emotion, 10*, 159-168.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotions in the human face.* Elmsford, NY: Pergamon Press.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., Tomita, M., & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology, 53*, 712-717.

Ekman, P., & Heider, K. (1988). The universality of a contempt expression: A replication. *Motivation and Emotion, 12*, 303-308.

Ekman, P., & Irwin, W. B. (1994). *An emotion lexicon for analyzing free responses in the judgment of facial expression.* Unpublished manuscript.

Ekman, P., & Oster, H. (1979). Facial expressions of emotion. *Annual Review of Psychology, 30*, 527-554.

Ekman, P., O'Sullivan, M., & Matsumoto, D. (1991a). Confusions about context in the judgment of facial expression: A reply to "The contempt expression and the relativity thesis." *Motivation and Emotion, 15*, 169-176.

Ekman, P., O'Sullivan, M., & Matsumoto, D. (1991b). Contradictions in the study of contempt: What's it all about? Reply to Russell. *Motivation and Emotion, 15*, 293-296.

Ekman, P., Sorensen, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science, 164*, 86-88.

Ellsworth, P. C., & Smith, C. A. (1988). From appraisal to emotion: Differences among unpleasant feelings. *Motivation and Emotion, 12*, 271-302.

Friesen, W. V., & Ekman, P. (1992). *Changes in FACS scoring.* San Francisco: Human Interaction Laboratory, University of California.

Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York: McGraw-Hill.

Holder, H. B., & Kirkpatrick, S. W. (1991). Interpretations of emotion from facial expressions in children with and without learning disabilities. *Journal of Learning Disabilities, 24,* 170-177.

Izard, C. E. (1971). *The face of emotion.* New York: Appleton-Century-Crofts.

Izard, C. E. (1977). *Human emotions.* New York: Plenum Press.

Izard, C. E., & Haynes, O. M. (1988). On the form and universality of the contempt expression: A challenge to Ekman & Friesen's claim of discovery. *Motivation and Emotion, 12,* 1-16.

Lazarus, R. S. (1991). *Emotion and adaptation.* New York: Oxford University Press.

Levy, R. I. (1973). *Tahitians: Mind and experience in the Society Islands.* Chicago: University of Chicago Press.

Levy, R. I. (1978). Tahitian gentleness and redundant controls. In A. Montague (Ed.), *Learning non-aggression.* New York: Oxford University Press.

Levy, R. I. (1984). Emotion, knowing, and culture. In R. A. Shweder & R. A. LeVine (Eds.), *Culture theory: Essays on mind, self, and emotion.* Cambridge, England: Cambridge University Press.

Matsumoto, D. (1992). More evidence for the universality of a contempt expression. *Motivation and Emotion, 16,* 363-368.

Matsumoto, D., & Ekman, P. (1988). *Japanese and Caucasian facial expressions of emotion (JACFEE).* Slide set and brochure available from first author, San Francisco State University, Department of Psychology, 1600 Holloway Ave. San Francisco, CA 94132.

Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation & Emotion, 13,* 143–157.

Ricci-Bitti, P. E., Brighetti, G., Garotti, P. L., & Boggi-Cavallo, P. (1989). Is contempt expressed by pancultural facial movements? In J. P. Forgas & J. M. Innes (Eds.), *Recent advances in social psychology: An international perspective* (pp. 329-339). Amsterdam: Elsevier.

Rosenberg, E. L., & Ekman, P. (1994). Coherence between expressive and experiential systems in emotion. *Cognition & Emotion, 8,* 201-229.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.

Russell, J. A. (1991a). The contempt expression and the relativity thesis. *Motivation and Emotion, 15,* 149-168.

Russell, J. A. (1991b). Rejoinder to Ekman, O'Sullivan, and Matsumoto. *Motivation and Emotion, 15,* 177-184.

Russell, J. A. (1991c). Negative results on a reported facial expression of contempt. *Motivation and Emotion, 15,* 281-291.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions: A review of cross-cultural studies. *Psychological Bulletin, 115,* 102-141.

Russell, J. A., & Fehr, B. (1987). Relativity in the perception of emotion in facial expressions. *Journal of Experimental Psychology: General, 116,* 223-237.

Russell, J. A., & Fehr, B. (1994). Fuzzy concepts in a fuzzy hierarchy: Varieties of anger. *Journal of Personality and Social Psychology, 67,* 186-205.

Scherer, K. R., & Wallbott, H. G. (Ed.). (1986). *Experiencing emotion: A cross-cultural study.* Cambridge, England: Cambridge University Press.

Shott, S. (1990). *Statistics for health professionals.* Philadelphia: W. B. Saunders.

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.

Tooby, J., & Cosmides, L. (1987). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology & Sociobiology, 11,* 375-424.

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior, 17,* 3-28.

Weiner, B. (1985). An attributional theory of achievement and motivation. *Psychological Review, 92,* 548-573.

West, M. (1953). *A general service list of English words.* New York: Longmans.

Wierzbicka, A. (1986). Human emotions: Universal or culture-specific? *American Anthropologist, 88,* 584-594.