Criteria for Emotion Recognition From Verbal and Nonverbal Expression: Studying Baggage Loss in the Airport

Klaus R. Scherer
Grazia Ceschi
University of Geneva

The issue of emotion recognition in real-life settings, lacking a clear criterion of the nature of the underlying emotion, is raised. After reporting their luggage lost, 110 airline passengers were asked to rate their emotional state (subjective feeling criterion). The agents who had processed the claims were asked to rate the passengers' emotional state (objective behavior criterion) as well as their own feelings. An excerpt of the videotaped interaction for 40 passengers was rated for emotional state by judges on the basis of (a) verbal and nonverbal cues or (b) nonverbal cues only. As predicted, the data show that judges' inferences in both exposure conditions correlate more strongly with the objective behavior (agent ratings) than the subjective feeling criterion (self-ratings). Using the Facial Action Coding System (FACS), objectively coded "felt" (but not false) smiles correlated positively with a good humor scale in both criteria and judges' ratings.

An important issue that often is neglected in the psychology of emotion concerns the nature of the criterion to be used in studies on the recognition of the underlying emotion from outward expression. Emotion is frequently defined as a psychological construct consisting of several components, including subjective feeling, motor expression, physiological arousal, and motivational action tendencies (Frijda, 1986; Lazarus, 1991; Scherer, 1984). The subjective feeling state component can be conceptualized as a reflection of the changes in all other components (Scherer, 1984, 1993). In consequence, we can assume that a person's awareness of a particular feeling is the result of a self-attribution based on auto-observation of the other components, that is, physiological symptoms, motor expression, and action tendencies.

In studies using posing or portrayal of emotion by actors (which constitute the bulk of the work in this area; see Banse & Scherer, 1996; Ekman, Friesen, & Ellsworth, 1972/1982; Gosselin, Kirouac, & Doré, 1995; Izard, 1971; Pittam & Scherer, 1993; Scherer, 1979), these different components are likely to be highly congruent. An actor is expected to base the expression on the appropriate antecedent appraisal of a given scenario, the associated physiological arousal and motor effector patterns, as well as felt action tendencies and feeling states (especially in posing based on Stanislavski techniques). In real-life emotions, this congruence may only exist for extremely intense emotions. In particular, it is possible that a person may simply not be aware of an emotional state that has produced a specific expression and/or that social control and regulation mechanisms prevent the accurate designation of the nature and intensity of the underlying emotional states.

Even if the person is aware of the ongoing changes in the other components, she or he might not be able to properly verbalize this feeling. Conscious representation and verbalization of emotional experience may only partially capture the reflection of synchronized changes in emotion components in a monitor system (see Kaiser & Scherer, 1997; Scherer, 1993; for a more detailed discus-

Authors' Note: The authors express their gratitude to the manager and staff of the baggage handling service of the international airline who granted field access and collaborated in the study. They acknowledge the contributions of Sabrina Scardua, Laurence Corthay Casot, Jaqueline Ritter, Leontine Berkhuijsen, Jean François Mibord, and Cordula Glimm who served as research assistants in the study, and of Ursula Scherer, who critically reviewed the manuscript. The authors particularly thank one anonymous reviewer for important data-analytical suggestions. Reprint requests should be directed to Klaus R. Scherer, Department of Psychology, University of Geneva, 40, bd. du Pont d'Arve, CH-1205, Geneva; e-mail: klaus.scherer@pse.unige.ch.

PSPB, Vol. 26 No. 3, March 2000 327-339

sion). Apart from the fact that verbalization can only reveal a small part of the complex changes occurring in an emotion episode, regulation and control mechanisms are likely to be operative at both conscious and unconscious levels, making it unlikely that verbal report is a faithful replica of underlying changes. In consequence, it is unlikely that a person's subjective verbal self-rating of his or her feeling state in a real-life emotion episode constitutes the sole valid criterion for all aspects of the underlying emotion process. Rather, we suggest that verbal self-ratings should be treated as constituting only one of several possible criteria, namely, a subjective feeling criterion.

Another possible criterion is the pattern of changes in the neurophysiological response systems. Psychophysiologists report evidence for "physiological emotions," clear patterns of physiological arousal that are not accompanied by subjective feelings (J. Fahrenberg, personal communication, June 9, 1998). Unfortunately, the current state of methodology does not easily allow researchers to obtain appropriate measures of this criterion in real-life settings.

Yet another potential criterion, which we shall call objective behavior criterion, might be based on motor expression and action tendencies. This criterion could be measured via observable behavior or inferred action tendencies as judged by persons who know the expresser intimately or by persons directly interacting with the expresser. It can be expected that judges having access to all expressive behavior and to details of the eliciting situation (which allow the reconstruction of the expresser's situational appraisal) will be able to produce a reliable judgment of both the expression and action tendency components of the expresser's emotion process. In addition, this criterion could be assessed by objective coding of the observable behavior.

In consequence, we suggest abandoning the notion of a single accuracy score in judges' emotion inference from expressive behavior in naturally occurring emotion episodes. Instead, we propose to study the correlation of judges' inferences with different emotion criteria, in particular a subjective feeling criterion as obtained from self-ratings and an objective behavior criterion as rated by interaction partners.

Unfortunately, the psychology of emotion suffers from a serious difficulty in finding opportunities to study emotional episodes in real life, particularly in a social context (Scherer & Wallbott, 1994; Wallbott & Scherer, 1989). Even if one can gain access, under special circumstances, to an individual's spontaneous emotional expression at a particular point in time, one can rarely generalize from a single case to the general process of emotion expression. One possible remedy is to identify typical social situations that reliably provoke relatively intense emotional states in many different persons and to exploit such events for research purposes.

This approach was chosen for the field study reported below. It exploits a class of natural events occurring with increasing frequency in airports across the world: Passengers waiting in vain for their baggage to appear in front of an empty conveyer belt and having to report the loss to the baggage tracing service. The study was conducted in collaboration with an international airline that provided access to the baggage delivery part of a major airport and arranged for the cooperation of the office staff responsible for handling claims for and tracing lost baggage. The situation studied has high ecological validity for the expression of everyday emotions and thus satisfies some of the requirements for a realistic emotion-eliciting situation.

In this study, we examine whether it is feasible to evaluate the accuracy of observer-judges' emotion inferences with respect to two different but complementary criteria for the nature of the underlying emotion process. Specifically, the following questions were investigated:

> How well does the self-reported feeling state of a person (subjective feeling criterion) correspond to the other-attribution (objective behavior criterion) that an airline agent forms in the process of a naturally occurring interaction?

The situation studied here provides the possibility of investigating the relationship between the two emotion criteria suggested above in a realistic field setting that is highly comparable across different persons. Contrary to standard emotional expression recognition paradigms, this situation provides more or less controlled expressions of spontaneously felt affective states in many different communication channels. The answer to this first question will allow a first assessment of the degree of overlap between these two different indicators of the type of emotion experienced in a relatively standard emotion-eliciting situation.

We hypothesize that the two criteria are unlikely to correlate highly with each other in this particular situation because the objective behavior criterion reflects only some of the components of the emotion process—mostly outward expression—reflected in subjective feeling state. In addition, it is likely that at least some passengers, when questioned about their affective state, will try—consciously or unconsciously—to describe their feelings in a way that differs from their true experiences. This could be either in the service of appearing unruffled by the luggage loss (projecting a stoic, stiff-upper-lip type of attitude) or, conversely, exag-

gerating their irritation in line with popular expectations concerning this situation. More important, even if passengers report their feeling states honestly to external interviewers, they may have controlled their verbal utterances and expressive behavior in the interaction with the agent. Because these factors are likely to affect individuals differentially, we expect the correlation profiles between the two criteria across different traits to vary extensively among passengers.

> To what extent do the two emotion criteria correlate with emotion inferences by judges who have only limited access (both in terms of a given time slice and available channels) to motor expression and behavioral manifestations of action tendencies?

In this research, we are particularly interested in determining to what extent cues to underlying emotions are available in brief samples of verbal and nonverbal behavior. Recent work by Ambady and Rosenthal (1992, 1993) has shown that very brief samples of nonverbal behavior, as short as half a minute, permit observers to gain an impression of a person's affective, interpersonal attitudes that correlates rather highly with criteria based on objective indexes or judgments based on long-term observations. Although rather brief exposure to nonverbal cues might be sufficient in standard expression-recognition paradigms or in situations in which little control or regulation is expected, the present situation might differ in the sense that passengers may attempt to control their expression for the sake of politeness or other strategic concerns, thus rendering the inference of their true underlying feeling state more difficult.

In this study, we attempt to examine to what extent it is possible to judge another person's subjective feeling state on the basis of limited access to the affective signaling of that person by presenting a brief excerpt from a videotape of the person's behavior in the interaction. Whereas one group of judges is presented with a brief time slice of both the verbal and nonverbal (facial and vocal) behavior, the other group is exposed to facial and vocal (low-pass filtered speech) nonverbal cues only. This aspect of the experiment will give some indication as to the role of verbal cues in emotion communication.

First, we hypothesize that judges' ratings will show considerably higher correlations with the objective behavior criterion than with the subjective feeling criterion because they have access (although limited in time and extent) to the same components (expression and action tendencies) as the interaction partners. Second, we hypothesize that the ratings of judges having access to verbal in addition to nonverbal cues should show higher correlations with the objective behavior criterion on one hand (because they have access to verbally described action tendencies) and with the subjective feeling criterion on the other hand (because—at least in some cases—passengers might reveal their feeling state verbally).

METHOD

Passengers

The study included 112 passengers arriving at a major international airport who were unable to retrieve their baggage and consequently had to go to the adjoining baggage handling office; these passengers were videotaped and interviewed. Of the passengers, 60% were male and 40% were female; 12.7% were between 20 and 30 years old, 84.6% were between 30 and 60 years old, and 2.7% were older than 60 years; and 30.9% came from French-speaking countries, 29.1% came from English-speaking countries, 16.4% came from German-speaking countries, 8.2% were Asian, and 15.4% came from other countries.

Agents

To ensure anonymity and confidentiality as well as to avoid evaluation apprehension, the identity of the agents handling the passengers was not monitored. Thus, no details on their number, gender, age, or the number of passengers dealt with were collected. Approximately 12 airline employees, predominantly female, all fluent in English, French, and German, participated in the study.

Recording Procedure

Two female research assistants were stationed at a conveyor belt, holding on to a luggage cart and ostensibly waiting for baggage to arrive. A camera and video recorder, both battery powered, were hidden under some pieces of luggage on the cart. By moving and pointing the cart, the angle of filming could be controlled. After spotting a single passenger standing at an empty conveyor belt and obviously waiting for their luggage to arrive, they videotaped his or her behavior (the region of the face and the upper torso).

When the passenger had gone to the baggage claim office, two other research assistants stationed in the baggage claim office took over. One of the desks in the claims office had been prepared for surreptitious recording of the passenger's interactive behavior. A camera hidden behind the desk allowed videotaping of the full face and part of the upper body of the passenger. Two microphones, one for the passenger and one for the agent, were attached under the desk and out of the passenger's sight. When the passenger entered the office, one research assistant directed him or her to this desk.

The agents working at this prepared desk knew about the recording activity (and that they had to fill out a questionnaire after each interaction) but were unaware of the exact purpose of the study. They were instructed to interact in a routine fashion with the passengers directed to this desk.

The routine claims procedure, which was recorded in its entirety, generally consisted of the following phases:

1. greeting and obtaining information from the passenger about the ticket and the final destination;
2. describing the luggage, its type, and its contents;
3. ascertaining the forwarding address in case the luggage was found; and
4. handing the passenger a copy of the retrieval forms and saying good-bye.

In general, these interactions lasted about 10 minutes, in a few exceptional cases up to 20 minutes, and were conducted in French or English.

Passenger Interview

None of the passengers was aware that his or her behavior was being recorded. After a passenger had left the desk, he or she was approached by a research assistant who explained about the videotaping and asked him or her to answer a number of questions. Every passenger was given the choice to have the tape erased. All 112 passengers agreed to have the tape used for research purposes and all except 2 answered all the questions in the questionnaire, the two refusals being due to time constraints of the passengers.

The passenger questionnaire, of which French, English, German, and Italian versions had been prepared, was administered by a research assistant who spoke the preferred language of the passenger. Five major areas were addressed: nature of the travel, appraisal of the baggage loss, subjective feeling state, evaluation of the agent and the claim service, and personal background. The procedure and the results concerning the appraisal questions are reported elsewhere (Scherer & Ceschi, 1997).

Subjective feeling state was measured both before and after having interacted with the agent. The questions were as follows: "How did you feel when you realized that your baggage would probably not be delivered with that of the other passengers, before you went up to the baggage retrieval service?" and "How do you feel now?" In response to these questions, passengers rated their feeling state on 7-point scales (where 0 = *not at all* and 6 = *very much*) with respect to each of five emotion categories: angry/irritated, resigned/sad, indifferent, worried/stressed, in good humor.

Information about nationality, age, and sex was obtained at the end of the interview.

Agent Ratings

The claims agents filled out a questionnaire after each interaction. They rated the passengers' emotions for the time at the beginning (before-condition) and at the end of the interaction (after-condition) on the same emotion scales used in the passenger questionnaire (see above). In addition, they rated their own subjective feeling state during the interaction (before-condition) and after the interaction (after-condition), that is, while filling out the questionnaire, on the same scales.

Judgment Study

As in the studies reported by Ambady and Rosenthal (1992, 1993), brief samples of the passengers' behavior were prepared. The first minute of each interaction was used as a stimulus segment to ensure the presence of an unabated level of emotional arousal elicited by the baggage loss and comparability of the interaction content across passengers.[1]

From the total of 112 passengers, a subsample of 40 was selected using the following criteria: acceptable quality of the videotaped material (some videotapes were excluded because of excessive movement of the passenger, equipment malfunction, or other technical factors) as well as approximately equal representation of gender of passenger and of passengers from French, English, and other language areas. The *t* tests between these 40 cases and the 70 nonchosen cases did not show any significant differences in self-ratings of emotion.

These 40 1-minute segments were copied onto two judgment tapes, each containing 20 segments. The video channel showed the face and the upper body of the each passenger. Audio channel 1 contained the complete speech signal. In audio channel 2, the speech signal was electronically filtered (low-pass filtering at 400 Hz for female and at 300 Hz for male voices). This procedure removes speech intelligibility and voice quality (timbre) information while leaving most paralinguistic vocal cues intact (see Scherer, Feldstein, Bond, & Rosenthal, 1985). By selecting the appropriate audio channel, the judgment tapes were used in two rating conditions: (a) full audio, in which judges could make use of all nonverbal and verbal cues, and (b) filtered speech, in which judges had to rely on nonverbal cues exclusively.

Students in an introductory psychology course at the University of Geneva served as judges and received course credit for their participation. Thirty-one judges (22.6% male, 77.4% female; M = 22.5 years, SD = 4.1 years) took part in the full audio condition and 36 judges (36.6% male, 63.4% female; M = 23.3 years, SD = 4.5 years) took part in the filtered speech condition. The procedure was the same in both conditions except for the selection of the appropriate audio channel. Judges, in groups numbering six to nine persons, were seated in

two semicircles in front of a 21-inch video monitor and rated the emotions experienced by each passenger immediately after viewing each interaction segment. The tape was stopped and judges had 30 seconds for their judgments. The same five emotion-rating scales as in the passenger and the agent questionnaires were used. The order of the presentation of two stimulus tapes was systematically varied for each of the four groups of judges.

Coding of Passengers' Facial Expressions

To gauge the distal facial expression cues available to judges in the rating task, the Facial Action Coding System (FACS) (Ekman & Friesen, 1978) was used. A preliminary screening of the judgment tapes had shown that the occurrence of expressions other than smiles was very low. The coding was therefore restricted to the facial action units that are the most representative for "felt" smiles (also often referred to as "enjoyment smile" or "Duchenne smile"), that is, Action Units 6 and 12 occurring in combination, and "false" smiles (nonfelt, polite smiles), that is, Action Unit 12 in isolation (see Ekman & Friesen, 1982; Frank & Ekman, 1993; Frank, Ekman, & Friesen, 1993).

An experienced FACS coder determined the number of episodes or instances of these two types of smiles as well as their respective durations for the 40 video segments. Because the overall reliability of the FACS system has been repeatedly demonstrated (see studies reviewed in Rosenberg & Ekman, 1998); the coder had passed the final examination administered by Ekman's laboratory, establishing her personal reliability; and the two action units coded here are among the most clearly discernible units, no further reliability check was computed.

An acoustic analysis of the vocal cues in the passengers' verbal utterances could not be performed because the quality of the recordings, done unobtrusively in a large reception area with several service desks, did not permit extracting the relevant parameters.

RESULTS

The passengers' self-ratings, the agents' attribution of passenger state, and the agents' self-ratings had been obtained for two time points—before and after the interaction (both collected retrospectively after the interaction). Because these highly correlated measures cannot be considered to be independent, only the ratings for the subjective feeling state before or at the beginning of the interaction will be used in the data analyses below.[2]

Relationships Between the Two Emotion Criteria

Passenger self-rating reliability. Before looking at the correlations of passenger self-ratings with the agent attributions, it is useful to examine to what extent passengers have a common emotional reaction to the luggage-loss experience using the profile of the ratings across the five scales. One way to examine this issue is to compute the intraclass r (the average r per passenger) and the Spearman-Brown "upped" equivalent (r of the entire set of passengers) using a repeated-measures ANOVA of the five scales × 110 passenger matrix to estimate the values (see Rosenthal, 1987, pp. 13-17).[3] In this case, we find intraclass $r = .009$ and effective $R = .509$. The very low reliability indicates the existence of a wide variety of different reaction patterns exhibited by our passengers. Thus, the claims of appraisal theorists that objectively highly similar situations can yield widely varying emotional reactions due to the differences in subjective appraisal of the situation are confirmed (see Scherer & Ceschi, 1997). Methodologically, it is important to have sufficient variance in the passenger self-ratings to yield valid correlations with agent attributions.

Correlations between passengers' self-ratings (subjective feeling criterion) and agents' attributions (objective behavior criterion). The Pearson product-moment correlations between passenger self-ratings and agent attributions for the five scales, shown in column 1 of Table 1, provide a first answer to Question 1 above.[4] The table shows significant and substantial positive correlations for anger/irritation, worry/stress, and good humor, whereas the association is at chance level for resignation/sadness and indifference. Scatterplots for the significant correlations revealed that the anger/irritation relationship is based on six extreme cases, raising doubts about the stability of this correlation.

Before we can discuss the significance of these findings, we need to deal with a number of possible artifacts. One issue consists of the potential effects of passenger background variables, such as gender, age, or country of origin, on the self-ratings of subjective feeling. To evaluate their importance, we regressed these three variables onto each of the five self-ratings, saved the residuals, and recomputed the correlations with agent attributions using the residuals (cf. column 2 of Table 1). Given that the correlation coefficients are very similar to those in column 1, it is unlikely that background variables had a strong influence on self-ratings of subjective feeling state.

Another possible artifact could be due to the agents' attributions being influenced, in the form of projection, by agents' self-ratings of their own feelings. These, in turn, may have been influenced by the passenger's affect state through a process of emotional contagion, empathy, or direct reaction (e.g., becoming angry due to the other's display of anger). In fact, a projection hypothesis seems plausible in light of the consistently high correlations between agents' self-ratings and other-attributions

TABLE 1: Intercorrelations of Passenger Self-Ratings (PSR), Agent Self-Ratings (ASR), and Agent Attributions (AA)

	Passenger Self-Rating/ Agent Attribution			*Passenger Self-Rating/ Agent Self-Rating*		*Agent Self-Rating/ Agent Attribution*
	Raw Passenger Self-Rating	*Residuals of Passenger Self-Rating*	*Agent Self-Ratings Partialled Out*	*Raw Agent Self-Ratings*	*Agent Attributions Partialled Out*	*Raw Agent Self-Ratings*
Angry/irritated	0.41***	0.40***	0.32***	0.27***	0.04	0.58***
Resigned/sad	0.04	0.06	0.15	–0.26***	–0.28***	0.34***
Indifferent	0.04	0.03	0.10	–0.07	–0.11	0.50***
Worried/stressed	0.33***	0.27***	0.32***	0.08	–0.03	0.35***
In good humor	0.40***	0.37***	0.33***	0.24*	0.04	0.53***

NOTE: $N = 110$.
*$p < .05$. ***$p < .001$.

(column 6 in Table 1). This suggests that these two sets of ratings are not independent of each other.

To examine the possibility of a projection of the agent's own feeling state on the feeling attributed to the passenger, partial *r*s between passenger self-rating and agent attribution, partialling out agent self-ratings, were computed (column 3 in Table 1). The results show that the magnitude of the partial *r*s is only very slightly different from the bivariate *r*s. Thus, the likelihood is very low that the correlations between passenger self-ratings and agent attributions are biased by agent self-ratings.

Another potential artifact might result from contagion or affect transfer, whereby the passengers' subjective feeling state might influence agents' emotions (measured via their self-ratings) and thus indirectly affect the agents' judgment of the behavioral emotion criterion (their attribution of passenger emotion). Column 4 in Table 1 does, in fact, show significant, although not very high, positive correlations between passenger and agent self-ratings for the anger and good humor scales. Thus, one might be tempted to assume a process of contagion and/or escalation in the case of anger and of simple contagion in the case of good humor.

To investigate this possibility, partial correlations between passenger self-ratings and agent self-ratings, controlling for agent attribution of passenger emotions, were run (see column 5 of Table 1). The two significant positive correlations drop nearly to 0. This suggests that little affect transfer or reactivity has occurred. The significant negative correlation for resignation/sadness remains. However, a scatterplot shows that 86.4% of the agents have values of 0 on this scale (which is understandable because, of the five emotions studied, it is the least appropriate for their role in the interaction). Thus, the correlation is strongly affected by restricted variance and should be interpreted with caution.

Having dealt with potential artifacts, judged to be of little or no consequence, we can return to the major result—the covariation between the subjective feeling criterion (as measured by passengers' self-ratings) and the objective behavior criterion (as measured by agents' attributions). Although the two criteria do overlap, the shared variance is actually quite low, hardly exceeding 15% for three of the scales and being close to 0 for two scales. Given the potential instability of the correlation for anger/irritation mentioned above, the two emotion criteria seem to have common variance only with respect to worry/stress and good humor. We will return to the significance of this finding for the issue of the relationships between emotion components in the Discussion section.

Accuracy of Emotion Recognition Based on a Brief Sample of Nonverbal Behavior

Judge reliabilities. As suggested by Rosenthal (1987, pp. 13-17) we computed the intraclass *r* (reliability of an average single judge) and the effective *R* (reliability of the mean judgment of all judges) for both groups of judges and all scales. The results are shown in Table 2. As one might expect, given the number of judges used, the effective *R*s are very high. In contrast, the reliability of a single average judge is relatively low, a phenomenon that is often observed in rating studies using lay observers. However, because we are not interested in judgments by single judges but rather in an estimate of the central tendency in the impressions of a typical group of observers, it is the size of the effective *R* that is most pertinent to our present purposes. Because this article does not focus on judge characteristics, we will not explore this issue further.[5]

Correlations Between Judges' Ratings and the Two Emotion Criteria

Question 2 above is addressed by the correlations between the judges' ratings and both passenger self-ratings (subjective feeling criterion) and agent attributions for the 40 cases, shown in Table 3. The ratings of the two judge groups correlate very highly with each

TABLE 2: Reliabilities of Judges' Ratings in Two Exposure Groups

	Judges' Ratings of Filtered Speech		*Judges' Ratings of Full Speech*	
	Effective R	*Intraclass* r	*Effective* R	*Intraclass* r
Angry/irritated	0.93	0.27	0.94	0.35
Resigned/sad	0.91	0.23	0.91	0.25
Indifferent	0.80	0.10	0.77	0.10
Worried/stressed	0.91	0.22	0.90	0.23
In good humor	0.98	0.55	0.96	0.46

NOTE: $N = 31/36$.

other, and *t* tests for differences between the two groups did not yield any significant differences. In consequence, the availability of verbal cues does not seem to have had a major effect on the correlation of judge ratings with either criterion.

The correlations of the judgments' subjective feeling criterion (passenger self-ratings) show, for both judgment conditions, only marginally significant relationships for one of the scales—good humor. In contrast, for the objective behavior criterion there are significant correlations for three of the scales—anger/irritation, worry/stress, and good humor. As predicted for the behavior criterion, agents and judges based their assessment at least in part on similar cues.

Table 3 shows some additional patterns. As expected, the anger/irritation correlation based on a few extreme cases disappears. Furthermore, most of the significant coefficients seem to be due to strong intercorrelations within the judge ratings rather than to genuine confusions. From these data, it appears that good humor is the only emotion scale where there is some degree of overlap between the subjective feeling and the behavioral criterion as well as an appreciable correlation with the judges' rating for both criteria. The question that needs to be raised is whether these correlations are due to the same persons being judged as low or high on good humor by all four groups (because it is possible that the pairwise correlations between any two groups are based on agreements with respect to different subgroups of passengers). A factor analysis of the four different types of good humor judgments yields only one factor (eigenvalue > 1), and all four types of ratings load positively on the factor (loadings: self-rating: .60, agent attribution: .68, full audio judges: .93, filtered speech judges: .89).

The only other scale that displays a tendency toward agreement across all four groups is worry/stress. However, here we find significant correlations between passengers and agents on one hand and between agents and the two judge groups on the other (there were no significant correlations between self- and judge ratings). As one might expect, in this case, a factor analysis of the four different types of worry/stress ratings yields at least two factors (eigenvalue > 1) with the following loadings: Factor 1: self-rating: .16, agent attribution: .38, full audio judges: .94, filtered speech judges: .97; and Factor 2: self-rating: .88, agent attribution: .76, full audio judges: .13, filtered speech judges: –.02. Here, clearly, the subgroups of passengers on which the correlations are based are different.

Correlations of the Ratings With Objectively Coded Felt and False Smiles

The FACS coding of smiles from the video segment showed a mean of .775 felt smiles ($SD = 1.165$) and an overall duration of 3.15 seconds ($SD = 5.328$). For false smiles, the video segment showed a mean of .350 ($SD = .735$) and an overall duration of 2.631 seconds ($SD = 2.631$). More than half of the 40 passengers studied showed a felt smile at least once, and in quite a few cases, more often than once. However, false, or polite, smiles, using only lip corner retraction, were used by only a quarter of the group, often only once. There were no significant differences in the frequency of these expressions with respect to gender, age group, or profile of the subjective appraisal variables (as described in Scherer & Ceschi, 1997). The two variables measured, number of episodes and overall duration of smiling, correlate highly (felt $r = .914$, false $r = .954$).

Both parameters for the two types of smiles were then correlated with passengers' self-ratings, agents' other attributions, and the ratings of the two judge groups. The results are shown in Table 4. Two-tailed tests of significance were used except in the case of the in good humor variable, where significant positive correlations had been expected on the basis of plausibility and findings reported in the literature.

These expectations are strongly confirmed: For all four types of ratings, there are significant positive correlations between both the number and duration of felt smiles and the good humor ratings. These correlations are extremely high, accounting for 60% to 70% of the variance for the two judge groups. This is understandable given the scarcity of information available to the judges. Their ratings of the good humor variable seem to be based exclusively on the extent of felt smiling. It is remarkable that false smiles seem not to have had any effect on the judgments given the consistently low correlations. Although this may be partly due to the lower incidence of this type of smiling and the reduced variance, the finding does suggest that human observers carefully distinguish between felt and false smiles in inferring the underlying affective state.

Correlations between agents' other attributions of good humor and felt smiles are lower, accounting for approximately 10% to 16% of the variance, but are still

TABLE 3: Intercorrelations Between Passenger Self-Ratings, Agent Other Attributions, and Judge Ratings in Two Exposure Conditions

	Passenger Self-Ratings					*Agent Attributions*				
	Angry/ Irritated	*Resigned/ Sad*	*Indifferent*	*Worried/ Stressed*	*In Good Humor*	*Angry/ Irritated*	*Resigned/ Sad*	*Indifferent*	*Worried/ Stressed*	*In Good Humor*
Passenger self-ratings										
Angry/irritated										
Resigned/sad	*0.31*									
Indifferent	*–0.37*	–0.26								
Worried/stressed	*0.50*	0.28	*–0.34*							
In good humor	*–0.58*	–0.23	0.47	*–0.38*						
Agent attributions										
Angry/irritated	–0.01	0.04	0.13	–0.02	0.19					
Resigned/sad	0.16	0.19	–0.14	0.17	*–0.36*	0.14				
Indifferent	–0.23	–0.08	0.06	–0.15	–0.07	*–0.33*	–0.05			
Worried/stressed	0.18	0.01	0.04	*0.39*	–0.09	*0.51*	0.29	–0.11		
In good humor	*–0.45*	–0.03	0.29	*–0.34*	*0.43*	–0.17	–0.01	0.25	–0.17	
Judges' ratings of filtered speech cues										
Angry/irritated	0.10	0.08	–0.12	–0.13	–0.11	*0.37*	–0.18	–0.03	0.21	–0.21
Resigned/sad	0.01	0.14	–0.08	0.12	–0.21	–0.08	–0.10	0.24	0.08	–0.13
Indifferent	–0.02	0.12	0.05	0.07	0.02	–0.28	0.03	0.03	*–0.32*	0.09
Worried/stressed	0.11	–0.08	–0.11	–0.14	–0.04	0.21	–0.05	0.14	0.30	–0.13
In good humor	–0.23	–0.17	0.17	–0.03	0.27	–0.21	0.10	–0.01	–0.23	*0.38*
Judges' ratings of full speech cues										
Angry/irritated	0.06	0.08	–0.07	–0.12	–0.13	*0.34*	–0.12	–0.06	0.29	–0.14
Resigned/sad	0.16	0.07	–0.17	0.26	*–0.34*	–0.05	0.17	0.25	0.20	–0.30
Indifferent	–0.18	0.03	0.15	–0.03	0.08	–0.22	0.02	–0.05	*–0.33*	0.09
Worried/stressed	0.24	–0.06	–0.13	0.04	–0.11	0.20	0.09	0.16	*0.34*	–0.21
In good humor	–0.29	–0.20	0.29	–0.12	*0.38*	–0.19	0.05	–0.02	–0.26	*0.45*

NOTE: $N = 40$, significant rs are set in italics, $p < .05$, two-tailed.

quite significant. Because the agents had much more information on each passenger's behavior and the overall context and significance of the baggage loss, they did not need to rely as much on this expressive pattern. Given the abundance of information available to the agents, it is not clear to what extent they used the smiles as a basis for their inferences. Again, the correlations with false smiles are very low.

In the case of the passenger self-ratings of good humor, the correlations with felt smiles drop further. They account for less than 10% of the variance and are only barely significant at the $p < .05$ level using a one-tailed test. A scatterplot shows that the low correlation is due to two groups of passengers: (a) those who do not smile at all but still report fairly high levels of good humor and (b) those who smile quite a bit but report very low levels of good humor. Using the definition of subjective feeling as a reflection of all other components of an emotional episode, as suggested in the Introduction, one can argue that felt smiles are reflected in the subjective experience, as verbally reported, but seem to be only one of many different factors that determine the nature of self-reported feeling. False, or polite, smiles are not at all reflected in the subjective feeling as reported, probably due to the fact that they are discounted as strategic behaviors in self- as well as in other-inference.

Although felt smiles correlate with good humor in the case of the self- and agent ratings, there are also highly significant negative correlations with the negative affect states in the case of the two judge groups. However, these are not independent phenomena. Most likely, judges decided on a good humor inference after observing frequent felt smiling and, consequently, in the interest of consistency, gave lower ratings for the negative emotions. The fact that these negative correlations are lower than the positive ones with good humor suggests that judges may still perceive some blending of emotions. Or, in other words, whereas a high rating of good humor excludes equally high ratings on negative emotions, there are cases where good humor is seen as accompanied by more negative emotions—as is the case for self-ratings (see Scherer & Ceschi, 1997, for an analysis of blending in self-reported affect states).

Overall, the consistently positive correlations between felt smiles and good humor ratings across all four types of emotion assessments explain why good humor is the only variable for which we find significant

Judges' Ratings of Filtered Speech					*Judges' Ratings of Full Speech*			
Angry/ Irritated	*Resigned/ Sad*	*Indifferent*	*Worried/ Stressed*	*In Good Humor*	*Angry/ Irritated*	*Resigned/ Sad*	*Indifferent*	*Worried/ Stressed*
0.19								
–0.53	0.07							
0.60	0.18	*–0.73*						
–0.66	*–0.65*	*0.35*	*–0.64*					
0.84	0.21	*–0.54*	*0.51*	*–0.53*				
0.11	*0.86*	0.01	0.27	*–0.62*	0.12			
–0.48	–0.05	*0.78*	*–0.72*	0.29	*–0.59*	–0.16		
0.51	0.19	*–0.68*	*0.87*	*–0.54*	*0.49*	*0.38*	*–0.77*	
–0.64	*–0.63*	*0.34*	*–0.58*	*0.95*	*–0.60*	*–0.64*	*0.38*	*–0.57*

correlations in all four groups. Given the results reported in Table 4, the shared variance of the ratings in all four groups can be clearly attributed to the frequency of felt smiles. This expressive behavior pattern is reflected to some degree in passengers' self-ratings, and it is one of the major, if not *the* major, cue used by external observers in inferring positive affect in a person.

DISCUSSION

We have argued that subjective feeling, even if reported correctly, will always reflect only parts of the overall changes in the different emotion components. One might expect the cognitive component, particularly the appraisal processes, and the motivational component, the action tendencies, to be overrepresented and the motor expression and physiological symptoms to be underrepresented. In addition, it is highly likely, particularly in a public situation such as the one studied here, that the conscious representation of feeling states, and in particular the verbal reports thereof, are affected by regulation and control attempts. We have suggested that a second, complementary criterion for the type of emotion should be obtained, via assessment of observable behaviors, to correct the suspected overemphasis on appraisal processes in the self-reported subjective feeling criterion. In the present study, agents' attributions of passenger emotions were used to estimate this criterion.

Question 1: Relationship Between the Two Emotion Criteria

We had hypothesized that the passengers' verbal reports of their subjective feeling state would not correlate very highly with the agents' attributions of the corresponding emotions because of the passengers' efforts to control the behavioral cues of their emotional state. The data in this study support this hypothesis. Significant correlations (in both the total sample and the subsample selected for the judgment study) are found for only two of the five emotions: worry/stress and good humor. These correlations account for about 15% of common variance. In the case of good humor, one can reasonably argue that the significant correlation between the two criteria is due to lack of control attempts. Thus, in a large intercultural study, Scherer and Wallbott (1994) report a significantly lower tendency to control or regulate the emotion in the case of joy as compared to negative emotions. In contrast, negative feeling states were controlled by most of the passengers, at least with respect to verbal

TABLE 4: Correlations Between Felt and False Smiles and Passenger Self-Ratings, Agent Other Attributions, and Judge Ratings

	Felt Smiles		*False Smiles*	
	Duration	*Number*	*Duration*	*Number*
Passenger self-ratings				
Angry/irritated	–0.22	–0.26	–0.14	–0.16
Resigned/sad	–0.24	–0.26	–0.22	–0.27
Indifferent	0.17	0.18	0.05	0.05
Worried/stressed	–0.13	–0.10	0.29	0.24
In good humor	0.16	0.27*	0.12	0.08
Agent attributions				
Angry/irritated	–0.22	–0.19	–0.14	–0.17
Resigned/sad	0.05	–0.01	–0.18	–0.17
Indifferent	0.10	0.08	–0.19	–0.12
Worried/stressed	–0.22	–0.24	0.14	0.11
In good humor	0.32*	0.40**	–0.13	–0.10
Judges' ratings of filtered speech				
Angry/irritated	–0.47**	–0.50**	–0.24	–0.27
Resigned/sad	–0.52**	–0.45**	–0.06	–0.06
Indifferent	0.31	0.31	0.02	0.02
Worried/stressed	–0.48**	–0.46**	–0.13	–0.13
In good humor	0.85**	0.84**	0.06	0.09
Judges' ratings of full speech				
Angry/irritated	–0.36*	–0.35*	–0.24	–0.29
Resigned/sad	–0.51**	–0.42**	–0.05	–0.08
Indifferent	0.19	0.15	0.10	0.11
Worried/stressed	–0.43**	–0.38*	–0.19	–0.20
In good humor	0.81**	0.79**	0.08	0.11

NOTE: $N = 40$.
*$p < .05$. **$p < .01$, two-tailed (except for in good humor = one-tailed).

utterances and expressive behaviors, if not necessarily in their verbal reports of feeling state to our interviewers.

Given the likelihood of such control attempts, how can we explain the fact that agents' attributions on worry/stress, a negative emotion, are significantly correlated with passengers' self-ratings? One possibility is that the agents, having learned about the significance of the luggage loss for each individual and the adverse consequences to be expected, have based their attributions on inferred appraisal processes rather than expressive cues. Alternatively, some of the passengers might have deliberately not controlled the expression of this emotion to generate empathy—and presumably greater effort to help—from the agent handling the baggage claim. Given the importance of control and regulation of emotion processes in social behavior, and the important strategic role the emotions play in interpersonal interaction (Fineman, 1996; Frank, 1993; Hochschild, 1993), it will be one of the major tasks for the study of naturally occurring emotions in real-life settings to disentangle the complex interactions between different propensities to control (or in some cases, accentuate) one's emotional expression depending on the perceived stakes.

Question 2: Relationship Between the Two Emotion Criteria and Judges' Inferences

We had hypothesized that judges' ratings would show considerably higher correlations with the objective behavior criterion than with the subjective feeling criterion because the judges have access to the same component cues (expression and action tendencies) as the agents but are not aware of the relative significance of the luggage loss to the respective passenger. The results confirm that this is indeed the case. The only trait for which there is a significant correlation with the self-ratings is good humor, whereas there are significant correlations with the agents' attributions for good humor, anger, and to some extent, also for worry/stress.

We have already commented on the fact that good humor is apparently less subject to regulation and control attempts (as demonstrated in other research referred to above) in explaining the large extent of overlap between self-ratings, agents' attributions, and judges' ratings. One may wonder why the judges' ratings share variance with the agents' attributions of anger and worry/stress but not resignation/sadness and indifference. There is a tendency (see Table 3) for judges to call resignation/sadness what agents describe as indifference. Similarly, passengers described by agents as worried/stressed tend to be rated as less indifferent by judges. Thus, it seems that the labels are used in a nonrandom and meaningful fashion but that the labels used for potentially similar behavioral expressions may be different.

Leaving the correlations with the self-ratings aside for the moment, one can ask how the effect sizes of the accuracy coefficients found in this study compare to those found in the classic emotion recognition studies in the facial and vocal domain. These coefficients have been obtained by using preselected, prototypical, high-intensity expressions, often produced by actors with the express purpose of being easily recognized, presented to judges in a highly focused, static fashion. Because facial expression studies have mostly used static photographs, and because the facial expressions in the present study (probably due to the pervasive control attempts) were relatively neutral, the data from studies on the vocal expression of emotion allow a more direct comparison. Comparing the present data with findings reported by Banse and Scherer (1996), we find correlation coefficients of very similar magnitude for anger (.34/.37 in this study vs. .34 in Banse and Scherer [1996]), worry or anxiety (.34/.30 vs. .42), and happiness (.45/.38 vs. .52).

Judging from this comparison, it would seem that the correlation coefficients obtained on the basis of brief slices of natural behavior in a task situation with rather controlled expressions, showing low intensity and little prototypicality, are rather impressive, underlining the judges' ability to recognize at least the objective behavior criterion for some central emotions with some, albeit not overwhelming, success. It might be argued that the judges in our study could use face, voice, and language cues for their inferences, as compared to the single modality stimuli generally used in classic recognition research (as mentioned above, facial cues are unlikely to have provided much information, with the exception of smiling). However, it is interesting to examine to what extent access to verbal cues improves judges' accuracy.

We had hypothesized that the ratings of judges who had access to verbal in addition to nonverbal cues should show higher correlations with the objective behavior criterion than should the judges who had access to nonverbal cues only. We also suggested that judges in the full audio condition might show slightly higher correlations with the subjective feeling criterion (or at least show similar profiles to the agents' attributions) because—at least in some cases—passengers might reveal their feeling state verbally.

Surprisingly, neither of these hypotheses was borne out by the data. The two judge groups show remarkably similar patterns (with respect to both size and configuration) of correlations with both self-ratings and agents' attributions, suggesting that the inferences were based on similar cues. It may well be that in the situation studied here, there is little room for anything other than strictly formal, informational exchanges. If this is the case, affective cues are carried mostly by the facial and paralinguistic channel. It should be noted that in the full audio condition, in addition to verbal material, judges had access to the upper frequencies of the voice, timbre, and voice quality (which is not the case for filtered speech). Again, this information does not seem particularly pertinent for the inference of the emotions studied here.

In addition to looking at mean accuracy, we calculated profile correlations across all five emotions for each individual passenger. On the whole, these profile correlations confirm what the earlier correlational analyses have shown: Although there is little accuracy with respect to the subjective feeling criterion, there is strong agreement with the behavioral criterion. This is true for both judge groups, yielding yet another indication that the availability of the verbal content adds little to the ability to infer the expressive component of the underlying emotion from nonverbal cues.

Nonverbal Cues Mediating Emotion Recognition

In trying to determine which cues judges used to infer emotions, particularly good humor, which was recognized rather well, we objectively coded smiling behavior. The results indicate that both agents' and judges' ratings of good humor strongly correlate with felt, but not false, smiles, indicating that these observers discriminated and made use of smiling cues.

At the same time, felt smiles, which are normally considered to be among the prime indicators of positive affect (Ekman & Friesen, 1982; Frank & Ekman, 1993; Frank et al., 1993; Izard, 1971), share only 10% of the variance with the self-ratings of good humor. This finding supports the assumption that verbally reported subjective feeling states are not necessarily a "gold standard" for the definition of the nature of an emotional episode unfolding in real life. This evaluation is not based on the difference in the size of the correlation coefficient between self-ratings and external judges' ratings. Clearly, if no other information is available, smiling—which is considered one of the most powerful mechanisms in the communication of affect—is likely to play a much more central role than it does for subjective feeling. The latter represents the integrated reflection of all emotion components, including the cognitive evaluation of the antecedent event and, most likely, the appraisal of the emotional reaction itself. These components are likely to be correlated among themselves rather than constituting completely independent determinants—otherwise emotional experience would be totally fragmented rather than integrated. In fact, it can be argued that the defining feature of emotion episodes is the coupling or synchronization of the component subsystems (Scherer, 1984, 1993; see also Rosenberg & Ekman, 1994, on coherence between facial expression and subjective feeling).

Furthermore, verbal self-reports of subjective feeling states are likely to be affected by social or self-control factors, as well as artifacts, in addition to changes in emotion components. Possible candidates for such artifacts—conscious or unconscious—are control or regulation mechanisms, social desirability, demand characteristics, social representations, and many other factors. In consequence, it seems quite risky to rely exclusively on self-report in trying to assess the nature of emotional experience.

Apart from providing promising insights with respect to the questions asked, the present study has shown that, although difficult, it is possible to study emotional phenomena under realistic conditions in the field. Regularly occurring events that are likely to generate fairly strong and variable emotions and are observable in public can

be profitably used in emotion research. Such semicontrolled field studies are urgently needed to overcome one of the major hurdles in emotion research, the difficulty of reliably (and ethically) producing reasonably strong emotions in the laboratory. It should be noted as a caveat, however, that control and regulation efforts may be particularly pronounced in public situations that are likely to provoke emotions. Future studies should address this issue more directly.

Most important, however, this study raises fundamental issues with respect to the definition of dynamic emotion processes in real-life settings and the need for multiple measurement criteria. Contrary to some of the more unidimensional approaches to emotion that have dominated the literature—emotion as valence, emotion as discrete neuromotor programs, emotions as subjective feelings—a closer look at the phenomenon in a realistic context confirms the need to treat emotions in a more fuzzy fashion, as bouts of constantly changing configurations of state changes in different components.

With respect to the classic issue of the recognition of emotion by external observers, this conceptualization raises the important issue of which criterion is to be used to define the nature of the underlying emotion. In line with a componential approach to emotion, we have argued that in contrast to classic recognition paradigms, multiple criteria need to be employed. The situation is further complicated by the fact that control and regulation attempts, depending on the situational context, may greatly affect the relative importance (and veridicality) of different components. In spite of its shortcomings, the present study may serve the essential function of reminding us of aspects of naturally occurring emotions that we must keep in mind if we aspire to an ecologically valid account of how emotions determine social interaction through the inference of others' emotional states.

NOTES

1. The video sequences filmed at the baggage conveyor belts were very diverse in nature (with respect to length, takes, technical quality, and other parameters) and thus did not afford the degree of comparability between stimuli that was deemed necessary for the judgment procedure.

2. As expected, most passengers reported blends or mixtures of different states (because they could indicate different intensities of several affect states in the interview). The details of the reported subjective feeling states and the nature of blending are discussed in greater detail in a companion article to the present report, in which the results of the appraisal data are reported (see Scherer & Ceschi, 1997).

3. The computation is as follows:

$$\text{intraclass } r\text{: } (\text{MSscales} - [\text{MSscales} \times \text{Passengers}]) / (\text{MSscales} + [(110 - 1)\ \text{MSscales} \times \text{Passengers}])$$

$$\text{Effective } R\text{: } (\text{MSscales} - \text{MSscales} \times \text{Passengers})/(\text{MSscales})$$

4. For economy of space and ease of comparison, Table 1 also shows all other types of correlations, to be discussed later in this article.

5. For readers interested in gender differences, we note that *t* tests between male and female judges did not yield consistent differences, with the exception of male judges in both groups attributing a higher level of resignation/sadness to passengers.

REFERENCES

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*, 256-274.

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology, 64*, 431-441.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*, 614-636.

Ekman, P., & Friesen, W. V. (1978). *The Facial Action Coding System*. Palo Alto, CA: Consulting Psychologist Press.

Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiles. *Journal of Nonverbal Behavior, 6*(4), 238-258.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1982). *Emotion in the human face: Guidelines for research and an integration of findings* (2nd ed.). Cambridge, UK: Cambridge University Press. (Original work published 1972)

Fineman, S. (1996). Emotions and organizing. In S. Clegg, C. Hardy, & W. Nord (Eds.), *Handbook of organization studies* (pp. 543-564). London: Sage Ltd.

Frank, M. G., & Ekman, P. (1993). Not all smiles are created equal: The differences between enjoyment and nonenjoyment smiles. Special Issue: Current issues in psychological humor research. *Humor International Journal of Humor Research, 6*(1), 9-26.

Frank, M. G., Ekman, P., & Friesen, W. V. (1993). Behavioral markers and recognizability of the smile of enjoyment. *Journal of Personality and Social Psychology, 64*(1), 83-93.

Frank, R. H. (1993). The strategic role of the emotions. *Rationality and Society, 5*, 160-194.

Frijda, N. H. (1986). *The emotions*. New York: Cambridge University Press.

Gosselin, P., Kirouac, G., & Doré, F. Y. (1995). Components and recognition of facial expression in the communication of emotion by actors. *Journal of Personality and Social Psychology, 68*, 1-14.

Hochschild, A. (1993). *The managed heart*. Berkeley: University of California Press.

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

Kaiser, S., & Scherer, K. R. (1997). Models of "normal" emotions applied to facial and vocal expressions in clinical disorders. In W. F. Flack, Jr., & J. D. Laird (Eds.), *Emotions in psychopathology* (pp. 81-98). New York: Oxford University Press.

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

Pittam, J., & Scherer, K. R. (1993). Vocal expression and communication of emotion. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 185-198). New York: Guilford.

Rosenberg, E. L., & Ekman, P. (1994). Coherence between expressive and experiential systems in emotion. *Cognition and Emotion, 8*(3), 201-229.

Rosenberg, E., & Ekman, P. (Eds.). (1998). *What facial expression reveals about emotion, development, psychopathology, and health*. New York: Cambridge University Press.

Rosenthal, R. (1987). *Judgment studies: Design, analysis, and meta-analysis*. Cambridge, UK: Cambridge University Press.

Scherer, K. R. (1979). Non-linguistic indicators of emotion and psychopathology. In C. E. Izard (Ed.), *Emotions in personality and psychopathology* (pp. 495-529). New York: Plenum.

Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Ed.), *Approaches to emotion* (pp. 293-317). Hillsdale, NJ: Lawrence Erlbaum.

Scherer, K. R. (1993). Neuroscience projections to current debates in emotion psychology. *Cognition and Emotion, 7*, 1-41.

Scherer, K. R., & Ceschi, G. (1997). Lost luggage: A field study of emotion-antecedent appraisal. *Motivation and Emotion, 21*, 211-235.

Scherer, K. R., Feldstein, S., Bond, R. N., & Rosenthal, R. (1985). Vocal cues to deception: A comparative channel approach. *Journal of Psycholinguistic Research, 14*, 409-425.

Scherer, K. R., & Wallbott, H. G. (1994). Evidence for universality and culture variation of differential emotion response patterning. *Journal of Personality and Social Psychology, 66*(2), 310-328.

Wallbott, H. G., & Scherer, K. R. (1989). Assessing emotion by questionnaire. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience. Vol. 4. The measurement of emotion* (pp. 55-82). New York: Academic Press.

Received November 19, 1997
Revision accepted December 31, 1998