Motivation and Emotion, Vol. 15, No. 2, 1991

Confusions About Context in the Judgment of Facial Expression: A Reply to "The Contempt Expression and the Relativity Thesis"[1]

Paul Ekman[2]
University of California, San Francisco

Maureen O'Sullivan
University of San Francisco

David Matsumoto
San Francisco State University

A recent study of the effect of context in the judgment of contempt facial expression (Russell, 1991) was flawed by several confusions about what constitutes context. We argue that the context used should have ecological validity, through the use of many, rather than a few, facial expressions, which are spontaneous rather than posed, and which are judged by carefully selected judgment tasks, using clearly defined or well-understood emotional terms. The confusion in Russell's work between accuracy studies and agreement studies is also addressed.

We agree with Russell (1991) that "[t]he face is judged not in an absolute manner but relative to the context of judgment," but we believe Russell's method — in particular his use of still photographs — is not appropriate to the study of such complex phenomena. When we used still

[1]Paul Ekman's work is supported by a Research Scientist Award from the National Institute of Mental Health (MH 06092). David Matsumoto's research is supported by a grant from the National Institute of Mental Health (MH 42749).
[2]Address all correspondence concerning this article to Paul Ekman, Human Interaction Laboratory, University of California, 401 Parnassus, San Francisco, California 94143.

0146-7239/91/0600-0169$06.50/0

photographs of facial expressions it was to answer simpler questions, such as whether people in different cultures will attach the same label to a given facial configuration. To examine how context influences the judgment of faces, which is Russell's intent, we believe the face should be seen as it usually is, in motion, competing, as it typically does, with other sources of information such as the body, voice, and speech. Using such methods, we (Ekman, Friesen, O'Sullivan, & Scherer, 1980) showed that the relative importance of face, body, speech content, and vocal cues in impression formation depends upon the interpersonal context. Others who also utilized video or audiotape (viz, Berry, 1991; Bugental, 1986; Krauss, Apple, Morency, Wenzel, & Winton, 1981) obtained similar results.

Even if one were to grant the utility of still photographs in studying the effects of context on judgment, we question the ecological validity (Brunswik, 1956) of how Russell is using such stimuli. Rather than examining how facial expression judgments are influenced by the context of many other expressions which the observer sees, Russell examined the effect of only one prior expression (what he calls the anchor) on one other photograph (the target). It is only with such minimal context that he is able to show the effect of context. When the context was multiple expressions (as in his Experiment 3), Russell obtained the same results we have reported. We suggest that our procedure — in which people see many expressions, not just two — is more true-to-life than Russell's procedure of showing just two expressions.

A more adequate test of Russell's ideas about contempt would show observers many different facial expressions, of many emotions, and within that context examine how different contempt expressions are judged depending upon what emotion is shown in the expression which immediately preceded each contempt expression. This paradigm was used by Matsumoto (1990), who had 42 observers view 99 photographs intended to depict different facial expressions. The observers used two judgment tasks: The first time they saw each expression they chose one emotion from a list of seven (anger, fear, disgust, contempt, surprise, sadness, happiness); the second time they saw the expressions they rated each picture on an intensity scale. Those data were collected and analyzed for a study of display rules.

To deal with the questions about context raised by Russell, we reanalyzed that data, focusing now on whether the emotion shown in the photographs which preceded contempt influenced the judgment of the contempt expressions.

In this data set photographs intended to depict contempt were preceded by photographs intended to depict anger three times, by disgust photos three times, by happiness photos three times, by fear photos twice,

Table I.

Contempt photograph preceded by photograph showing (anchor)	Percent judging contempt expression as contempt (target)	
Anger photo 1	Contempt photo 1	57.1
Anger photo 2	Contempt photo 2	52.4
Anger photo 3	Contempt photo 3	57.1
Disgust photo 1	Contempt photo 4	59.5
Disgust photo 2	Contempt photo 5	57.1
Disgust photo 3	Contempt photo 6	76.2
Fear photo 1	Contempt photo 7	66.7
Fear photo 2	Contempt photo 8	57.1
Happiness photo 1	Contempt photo 9	71.4
Happiness photo 2	Contempt photo 10	66.7
Happiness photo 3	Contempt photo 11	64.3
Surprise photo	Contempt photo 12.	57.1

and by surprise once. As Table I shows, in every instance the modal judgment of the photograph intended to depict contempt was contempt. Thus, the proximal context (the facial expression shown in the immediately preceding photograph) had no effect on the modal judgments of these pictures, regardless of what expression was shown in the preceding "anchor" photograph.

We believe these results differ from what Russell reports because we have examined the issue more thoroughly. In our study, the photographs were seen in a more representative context (many expressions rather than two), so that the judgment of contempt was examined when many different emotions preceded it. The data in Table I, as well as our other research, differ from Russell's experiments in another way which may also contribute to the difference in our findings. In our study, observers made categorical judgments prior to making any other evaluation of the expressions; in Russell's Experiments 1 and 2, the categorical judgments were made after the observers made quantitative ratings of the same pictures. In recent work, O'Sullivan and Seyranian (1991) have shown that the order in which judgment tasks are performed influences the results obtained.

They had 63 subjects rate facial expressions of emotion using two formats: (1) a free choice from seven affect terms and (2) an intensity rating of each of the seven affect terms. When subjects rated facial expressions with the free-choice format first, on average, 80% of them used only a

single affect term to describe the facial expressions. When the subject made intensity judgments first, and then were asked to make a free choice among the seven affect terms, 62% of them chose only one term. Subjects who were "primed" to evaluate the photographs more carefully (by giving intensity judgments first) gave more elaborated responses than those who did not have such an expectation. The kind of judgment asked for is an important aspect of context which Russell has not acknowledged.

Nonetheless, Russell's finding in Experiments 1 and 2 that the modal judgment was *not* contempt creates a problem for us to explain. First let us note that in the control groups contempt was always the second most frequent judgment and those who did not call these pictures contempt did not attribute random emotions to the expressions. Fear, anger, happiness, surprise, and sadness were virtually never attributed to these faces. As would be predicted from the theoretical writings of Ekman and Friesen (1975), Izard (1971), and Tomkin (1963), disgust was the most frequently chosen label (unless disgust was the anchor).

Still, why did Russell's observers call the contempt photos disgust when ours called them contempt? In the one study in which Russell showed observers photographs of six other emotions (Experiment 3), the modal judgment was contempt. This study resembled our studies in which observers always saw many different faces, not just one photograph of one emotion. Perhaps people recognize contempt only when they are exposed to a variety of emotional expressions, as they typically are in real life, and as they were in the study reported in Table I. In Russell's Experiment 3, however, they could have chosen contempt because it was an unused category. That probably was not the reason, since in our studies the observers saw so many different expressions they could not have easily used an exclusion rule. But it is an empirical question, which a student of judgment context, such as Russell, should investigate.

Another aspect of context is the lexical abilities of the judges. Russell suggests that his observers were less educated than ours. Perhaps a college education is needed to understand the word "contempt", at least in English-speaking countries. If Russell believes education to be a factor in judging facial expressions of contempt, he should obtain data to demonstrate that point. We think it likely that the term for contempt may be more difficult to understand in English (perhaps too similar in sound and form to *content*) than it is in other languages. This possibility is consistent with Matsumoto's (1991) observation from his study of judgments of faces in the United States, Poland, India, Japan, and Hungary. Only in the United States have subjects asked for clarification of the meaning of the word *contempt.*

This point is related to the issue of our providing a definition of contempt in some of our studies. In literate cultures it has usually been our practice to have the observers read written instructions which explain the task (Ekman, et al., 1987). These instructions do not define the emotion terms, but simply list each emotion category with one word designations (anger, contempt, disgust, fear, etc.). Then we have asked whether there are any questions before we begin the experiment. The most frequent question in English-speaking groups has been about the meaning of contempt, although sometimes questions are asked about disgust. In response to such questions, we have defined contempt as "to feel morally superior, or better than another," and disgust as "to feel repulsed by something or someone." Note that we have not defined contempt or disgust by giving single word synonyms or related but different emotion terms, as did Izard and Haynes (1988). If we had done that, we would not have known which of the emotion terms the observers had in mind when judging a face.

It is certainly fair to say, however, that some of the observers did have the definitions listed above in mind. It would be useful to know whether Russell would obtain the same results if he were to give those definitions to his observers. We would expect similar results, but with higher agreement, than with the one-word terms. In studying judgment context, data should be collected about the influence of observers' having and not having definitions of each emotion term, or different definitions of each term.

Question might also be raised about whether contempt is really different in some way from other emotional states. Ekman and Friesen (1986) chose to study contempt precisely because they thought it was very different from other emotions — late in appearance developmentally, not reported in other primates, involving a moral judgment, and unilateral in its expression. They had expected that there would *not* be cross-cultural agreement about the judgment of this expression, but there was. Perhaps contempt is less well defined, less salient in the English language because English-speaking countries are more individualistic, less communally oriented than other cultures. The fact that an emotion is not well-defined linguistically, however, does not necessarily mean, as Russell suggests, that it does not have a universal expression. The anthropologist Robert Levy (1984) reported that sadness was not linguistically identified in Tahitian, although Tahitians did show sad expressions and behavior in response to loss. Just because contempt is different from other emotions, it is imperative that students of judgment context not limit their studies to this one emotion. Could the same results be found with fear, anger, happiness, and sadness? More research is needed to answer this question.

Another aspect of context which is not adequately addressed in Russell's present and previous research (Russell & Fehr, 1987) is the place of such relativity research in the history of similar research endeavors. The work of Schlosberg (1952, 1954) and Abelson and Sermat (1962) is surely seminal in this area. The issue of why that earlier version of a relativity approach to emotion judgments did not bear fruit should be considered and not merely overlooked (perhaps causing us to reinvent a square wheel.)

From our vantage point, Russell is not only mistaken in how to study context effects in the judgment of facial expression, but he has also confused the difference between an agreement and an accuracy study. In his first sentence, Russell says that he "began with the question of whether people are *accurate* [emphasis added] when they judge someone as feeling fear, anger or contempt. . . ." At the end of the paper he acknowledges that he is not really studying accuracy and implies that it is next to impossible to do so. We agree he is not studying the question of whether the face provides accurate information, but instead is asking what factors influence agreement among observers. While Russell's question is important, we disagree with his statement that accuracy is not the best question to ask, and with the implication that it is a near impossible question to address because there is no way to know what the true emotion is. Russell takes this position, we believe, because he does not understand what an accuracy study requires.

In their review of the literature on facial expression, Ekman, Friesen, and Ellsworth (1972) defined accuracy "to refer to 'correct' information of some nature being obtained, by some means, from facial behavior" (p. 16). They described two very different means of obtaining information from the face: measuring facial behavior itself or asking observers to make judgments about the faces they see. The findings from these two methods may not always converge. Electromyographic measurement (EMG), for example, has found differences between depressed and nondepressed (Schwartz, Fair, Salt, Mandel, & Klerman, 1976) persons by measuring facial activity which is not visible to observers. Russell has used observers' judgments, not facial measurement, to extract information, but he did not have any criteria to determine whether those judgments were correct. Such criteria could not be established with our photographs (which Russell has used), which were not made to study accuracy but agreement, and which, excepting the happiness photographs, were produced by telling the posers what muscles to move. Russell's (1991) claim that ". . . all we have today are various other judgments" fails to recognize the nature of the stimuli which he is using and is misleading in implying that accuracy cannot be studied.

When spontaneous facial expressions are studied, Ekman and his colleagues described four different accuracy criteria which can be used to determine whether the information extracted from the face is correct: antecedent events, concomitant behavior, consequent events, and consensus by a panel of experts. They recommended utilizing more than one criterion and examining carefully situations where discrepancies among criteria occurred. There is considerable evidence that the face can provide accurate information about positive vs. negative emotions under some conditions, and about some, but far from all, the distinctions we and others believe can be made among positive and among negative emotions (cf. Matsumoto, Ekman & Fridlund, 1991; Fridlund, Ekman, & Oster, 1987).

It is difficult, but not impossible, to perform accuracy studies. Many interesting questions remain about when the information provided by facial behavior is accurate and how precise such information may be. We would not claim that this is the "best" question to ask about facial expression or emotion, but it is certainly remains a timely and viable one, just as the question Russell has chosen to address — what influences observers' agreement about the emotion shown in a face — is another worthy topic for investigation.

In conclusion, let us note our agreement with Russell's (1991) careful statement that the judgment of the emotion conveyed by a facial expression ". . . may depend on many subtle details of the context of that judgment . . ." We certainly hope he now does the research needed to find out how strong these effects may be, for which emotions, at what intensity, and whether these effects survive when facial expression is embedded within the context of other interpersonal behavior, not studied only as an isolated, statistic, posed representation in a still photograph.

REFERENCES

Abelson, R. P., & Sermat, V. (1962). Multidimensional scaling of facial expressions. *Journal of Experimental Psychology, 63*(6), 546-554.

Berry, D. S. (1991). Accuracy in social perception: Contributions of facial and vocal information. *Journal of Personality and Social Psychology, 61,* 298-307.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Bugental, D. (1986). Unmasking the "polite smile": Situational and personal determinants of managed affect in adult-child interaction. *Personality and Social Psychology Bulletin, 12,* 7-16.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face.* Englewood Cliffs, NJ: Prentice-Hall.

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

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings.* New York: Pergamon Press.

Ekman, P., 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. R., 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., Friesen, W. V., O'Sullivan, M., & Scherer, K. (1980). Relative importance of face, body and speech in judgments of personality and affect. *Journal of Personality and Social Psychology, 38,* 270-277.

Fridlund, A. J., Ekman, P., & Oster, H. (1987). Facial expression of emotion. In A. Siegman & S. Feldstein (Eds.), *Nonverbal behavior and communication* (pp. 143-224). Hillsdale, NJ: Erlbaum.

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

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

Krauss, R. M., Apple, W., Morency, N., Wenzel, C., & Winton, W. (1981). Verbal, vocal, and visible factors in judgments of another's affect. *Journal of Personality and Social Psychology, 40,* 312-320.

Levy, R. I. (1984). The emotion in comparative perspective. In K. R. Scherer, & P. Ekman (Eds.), *Approaches to emotion.* Hillsdale, NJ: Erlbaum.

Matsumoto, D. (1990). Cultural similarities and differences in display rules. *Motivation and Emotion, 14,* 195-214.

Matsumoto, D. (1991). [Cross-cultural judgments of facial expression]. Unpublished data.

Matsumoto, D., Ekman, P., & Fridlund, A. (1991). Analyzing nonverbal behavior. In P. W. Dowrick (Ed.), *Practical guide to use video in the behavioral sciences* (pp. 153-165). New York: John Wiley & Sons.

O'Sullivan, M., & Seyranian, G. (1991). [The effect of response format on ratings of facial expressions of emotion]. Unpublished raw data.

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

Russell, J. A., & Fehr, B. (1987). Relativity in the perception of emotion in facial expression. *Journal of Experimental Psychology, 116,* 233-237.

Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology, 44,* 229-237.

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review, 61,* 81-88.

Schwartz, G. E., Fair, P. L., Salt, P., Mandel, M. R., & Klerman, G. L. (1976). Facial muscle patterning to affective imagery in depressed and nondepressed subjects. *Science, 192,* 489-491.

Tomkins, S. S. (1963). *Affect, imagery, consciousness: Vol. 2. The negative affects.* New York: Springer.