

Methodological Problems in Tourangeau and Ellsworth's Study of Facial Expression and Experience of Emotion

Joseph C. Hager and Paul Ekman
University of California, San Francisco

Flaws in the method of Tourangeau and Ellsworth's study undermined their test of hypotheses derived from theories of emotion.

Tourangeau and Ellsworth (1979) attempted to test models of emotion that emphasize facial expression (see Izard, 1981; Tomkins, 1981) with an experiment "in which facial expression was manipulated independently of emotional stimulation" (p. 1522). To manipulate emotional stimulation, subjects watched either a sad, fear-arousing, or emotionally neutral film for 2 minutes. To manipulate facial expression, subjects were instructed to make and hold throughout the film facial movements corresponding to a fearful expression, a sad expression, or a grimace unrelated to emotion. Other subjects received no expression instructions. To prevent subjects from realizing the importance of the facial movements and their relation to emotion, the authors used a cover story. They gathered self-reports of emotion and observed facial expression and autonomic responses in each factorial combination of the film and expression manipulations. Thus, the authors examined responses when subjects' manipulated expressions were appropriate, inappropriate, or irrelevant to the emotional content of the film and when no expression was requested.

Three hypotheses were tested with this design. A "sufficiency" hypothesis that "an expression is sufficient to produce the experience" (p. 1521) would be disconfirmed if subjects who made an expression reported either no emotion or one inconsistent with the expression. A "necessity" hypothesis that "an appropriate facial expression is necessary for the subjective experience" (p. 1521) would be disconfirmed if subjects re-

ported an emotion without an appropriate expression. Finally, a "monotonicity" hypothesis predicted that a measure of the strength of expression should be positively correlated with subjects' self-reports of emotion.

Methodological flaws weakened the test of each hypothesis. The sufficiency hypothesis required subjects to make a valid analog of an emotion expression as a basis for experiencing emotion.¹ Subjects probably did not make valid analogs because (a) the facial movements requested were not as distinctive of sad versus fear expressions as the authors thought, (b) the movements were too difficult for all subjects to make, (c) the authors did not show that subjects made them, (d) the sad and fear expressions were so similar anatomically that slight errors in performance probably changed one into the other, and (e) the expressions may have been a better basis for physical discomfort than for the emotions intended.

A test of the necessity hypothesis required subjects to make only the requested expression to preclude other expressions as mediators of experience. However, (f) expressions besides the one requested could have occurred to mediate self-report, and (g) the authors failed to show that such expressions did not occur.

The monotonicity hypothesis required a method to distinguish different expressions and their strengths, but (h) the measure of

¹ Tomkins (1981) and Izard (1981) argue that voluntarily making movements frozen for 2 minutes in the absence of other necessary physiological activity does not produce valid analogs of expressions of emotion as described in their theories. We make a different point: that the movements were not valid even within Tourangeau and Ellsworth's own conceptual framework.

Requests for reprints should be sent to Paul Ekman, Department of Psychiatry, University of California, San Francisco, 401 Parnassus Avenue, San Francisco, California 94143.

expression was not sensitive enough to test this hypothesis.

Finally, the authors claimed that they avoided contamination of their measurements by experimental demand characteristics. Demand characteristics remain a viable alternative explanation because (i) subjects probably knew the meanings of requested expressions but were motivated to hide that they knew the experimenter lied, (j) the authors did not report a debriefing extensive enough to uncover demand characteristics created by their deceptive cover story, and (k) a demand to ignore the requested expression in favor of cues from the film when reporting emotion is a possible explanation of the authors' findings. Each of the above points is elaborated below.

The Movements Were Not Valid Analogs of Expressions of Emotion

(a) Tourangeau and Ellsworth said that they derived instructions for making expressions from Tomkins (1962), Izard (1971), and Ekman and Friesen (1975). The muscle actions and appearances requested for the sad and fear mouth expressions were not entirely consistent with the work of these authorities. The authors instructed subjects to make the fear mouth by "pulling the corners of your mouth down and back" (p. 1523). The authorities cited agree that lip corners back or lateral stretching, but not a downward movement, is the key element in an expression of fear in the mouth area. Lateral stretching is produced by the risorius muscle, which may incidentally produce either upward or downward angling of the lip corners. Instead of the risorius, the authors requested subjects to contract the triangularis, which is characteristic of a sad and not a fear expression according to Darwin (1872), Tomkins (1962), and Ekman and Friesen (1975). Izard is the only authority cited who has indicated that this muscle may be involved in expressions of fear. It is likely that the authors produced a blend of sad and fear movements instead of a pure fear expression because they specified triangularis contraction and emphasized lip corners down for the fear mouth condition. For the sad mouth condition, the authors wanted the "corners

of lips pulled down" (p. 1523), which is an action characteristic of sad expressions. To obtain this action, the authors requested subjects to contract the quadratus muscles (depressor labii inferioris), but this muscle lowers the center of the lip, not the corners. No authority cited has indicated that this muscle is important for sad expressions. The authors should have requested contraction of the triangularis, which pulls the corners of the lips down. The instruction to contract the quadratus might have misled subjects if any knew the anatomy of facial movements. Thus, the distinction between sad and fear expressions was blurred by confusing the crucial mouth movements for each and by mistaking their muscular basis. The distinction between these expressions would have been greater had the authors requested actions which the authorities all agree are important for these emotions (i.e., lip corners stretched laterally for fear). Alternatively, the authors could have contrasted emotions that have more distinct expressions (see point d).

(b) The brow and eye movements they requested were consistent with theory and research on the expressions of sadness and fear, but these movements are among the most difficult to make deliberately. Ekman, Roper, and Hager (1980) reported that few young children could make these combinations of movements and that only 4 of 12 children 13 years old could make them without using a mirror. Our unpublished testing of hundreds of adults on these movements has shown that most people cannot make them deliberately. This difficulty suggests that not all of the authors' subjects were able to make them.

(c) Since the requested movements were so difficult, hard evidence that subjects made them is necessary. The only assurance given was that an experimenter determined before starting the film that the movements were "approximately right" (p. 1523), which implies that exact performance was not required. Showing that trained observers could rate the expressions reliably on emotion scales is not evidence that the expressions were correct. As the authors stated, "raters might have recognized the intended expressions even though subjects' faces were poor

reflections of the canonical fear or sad expression" (p. 1528).

(d) Exact performance of the requested movements was needed to distinguish fear from sad expressions because there is considerable overlap in the anatomical basis and appearance of these expressions. Three of the four components in the sad expression were also in the fear expression.² Failure to contract the outer frontalis or to maintain this contraction would change the requested fear brow to the requested sad brow. Failure to contract the risorius or to maintain this contraction would change the requested fear mouth into a sad mouth. Changes from one expression to another were likely because, as the authors admitted, "subjects could not consistently maintain their instructed expression" (p. 1524). Even experts trained in moving their faces whom we asked to make these expressions could not hold them for as long as a minute. Expressions changed as some muscles relaxed or others contracted until the experts relaxed entirely. These changes tended to make the experts' expressions look more alike as time passed. By studying two emotions that have such similar expressions, the authors made their manipulation of expression vulnerable to slight errors in performance. This weakness could have been avoided if they had instead chosen emotions that have more contrasting expressions, such as disgust and fear.

(e) The facial experts who made the requested expressions for us complained of physical discomfort.³ The authors' instructions to subjects included the warning that the movements might be "somewhat uncomfortable" (p. 1522). The expressions may have been a better basis for discomfort than for any emotion. The theories of emotion the authors tested specify that such discomfort can elicit sadness, distress, fear, or other affects. Subjects' discomfort may have artifactually contributed to their experience of sadness or other affects in any condition.

Other Expressions Besides the Requested Movements Occurred

(f) Our observations under point *d* contradict the authors' view that making one deliberate expression would "greatly inter-

fere with the emergence" of the other (p. 1528). Likewise, the authors assumed that adopting a nonemotional expression prevented spontaneous expressions of emotion from occurring. As noted by Izard (1981), spontaneous expressions of emotion could have occurred when subjects failed to maintain the requested expression. They could also have blended with expressions subjects made deliberately. Such spontaneous expressions could have mediated subjects' self-reports of emotion consistent with the necessity hypothesis.

(g) The authors had no way to show that unrequested expressions did not occur. To assess subjects' facial behaviors, the authors used judges' ratings of facial expression on 9-point sad, unhappy, scared, and afraid scales. The reliability obtained for these ratings might have been due only to judges' ability to rate requested expressions. These ratings cannot be trusted to reflect accurately spontaneous, fleeting expressions that might have occurred in addition to the requested movements.

(h) Careful facial measurement (e.g., Ekman & Friesen's, 1978, Facial Action Coding System) could have determined exactly what muscle movements subjects made and established a reliable basis for inferring the strength of expressions. We agree with Tourangeau and Ellsworth that the condition in which subjects received no expression instructions was "especially relevant" (p. 1528) for looking at the relationship between expression and self-report. Their inability to find a relationship between expression and report is contradicted by Ekman, Friesen, and Ancoli's (1980) study, which happened to use one of the same stimulus films. However, Ekman et al. did not rely on judges'

² In terms of Ekman and Friesen's (1978) Facial Action Coding System, their instructions would probably be 1 + 2 + 4 + 5 + 15 + 20 and 1 + 4 + 15 + 17 for fear and sad, respectively, based on the authors' descriptions of the appearances they wanted to produce. This scoring reveals the overlap of components in these two expressions. 1 = inner frontalis; 2 = outer frontalis; 4 = corrugator; 5 = levator palpebrae superioris; 15 = triangularis; 17 = mentalis; 20 = risorius.

³ Izard (1981) reports that his physiological consultants predict fatigue from making these movements. This prediction is consistent with our experts' experience, since the fatigue becomes uncomfortable.

global ratings of emotion to score the face, but instead employed a fine-grained measurement technique that requires slowed and repeated inspection of the face. Perhaps Tourangeau and Ellsworth's judges were not skilled or sensitive enough to recognize such subtle facial signs of spontaneous emotion. Although Ellsworth is knowledgeable about facial measurement (e.g., she worked with Ekman and Friesen in developing their earlier and cruder facial measurement technique; Ekman, Friesen, & Tomkins, 1971), mere expertise cannot substitute for facial measurement. Our experience suggests that reliable discriminations between emotions as similar in appearance as fear and sadness cannot always be obtained through observation alone, even repeated observations by a skilled person. Certainty requires slowed motion viewing with precise identification of each of the muscular components shown in each expression.

Demand Characteristics Contaminated the Experiment

(i) The authors did not feel that demand characteristics had contaminated their experiment because "no one guessed that facial expression was the variable of interest, nor that his or her expression had been an emotional one" (p. 1523). They trusted subjects on this issue, but we do not. It is doubtful that Yale undergraduates were unable to realize the importance of the strange facial movements and their emotional meanings, especially when reporting their emotions. One of us read Tourangeau and Ellsworth's instructions for making expressions to undergraduates. They had no trouble understanding the appropriate meanings of these expressions on the four scales used in their study. If Tourangeau and Ellsworth's subjects became aware of the meanings of the requested expressions, they would have known that the cover story was false and suspected that the experiment was deceptive. The demand characteristic could have been not to reveal that they had stolen this hidden information. This might explain why "several subjects expressed confusion about the complex cover story" (p. 1523). Subjects might also have been motivated to dissim-

ulate because many were receiving class credit as well as cash for their participation.

(j) Given the possibility of such complex demand characteristics, an extensive debriefing procedure would have been necessary to ascertain the extent of subjects' insight (Orne, 1962). This debriefing would have given permission for subjects to admit that they knew the experimenter had lied. The authors did not report using such a comprehensive debriefing.

(k) Tourangeau and Ellsworth did not find any significant effects of the requested facial expression on self-report. A plausible explanation that cannot be ruled out by the authors' reported data is that subjects figured out that the true purpose of the experiment was to see whether they would rate their experience based on their expressions or on the film. Was there a demand characteristic that led subjects to ignore the meaning of the requested expressions in favor of cues elicited by the films when rating their emotional experience? The apparent artificiality and inappropriateness of the requested expressions might have led subjects to discount them as a basis for rating their experience. The films, however, elicited real experiences (e.g., genuine emotional expressions and physiological responses) that subjects could use to make their ratings.

The only other effect on self-report besides the films was an interaction between the experimenter who gave the instructions and the facial expression requested. The authors did not interpret this finding. A simple interpretation is that the two experimenters could not produce the same pattern of results.

Summary

Tourangeau and Ellsworth did not manipulate facial movements into valid analogs of emotional expressions, so failure to confirm the sufficiency hypothesis is uninformative. The manipulation of facial expression was not independent of the film manipulation because the films could have elicited spontaneous emotional expressions. Since the authors had no way to prevent or detect these expressions, they could not test their necessity hypothesis. Their monotonicity hypothesis required a more sensitive measure of the

strength of expression to provide an adequate test. Their results can be explained by demand characteristics.

References

- Darwin, C. *The expression of the emotions in man and animals*. London: Murray, 1872.
- Ekman, P., & Friesen, W. V. *Unmasking the face*. Englewood Cliffs, N.J.: Prentice-Hall, 1975.
- Ekman, P., & Friesen, W. V. *The facial action coding system*. Palo Alto, Calif.: Consulting Psychologists Press, 1978.
- Ekman, P., Friesen, W. V., & Ancoli, S. Facial signs of emotional experience. *Journal of Personality and Social Psychology*, 1980, 39, 1125-1134.
- Ekman, P., Friesen, W. V., & Tomkins, S. S. Facial affect scoring technique: A first validity study. *Semiotica*, 1971, 3, 37-58.
- Ekman, P., Roper, G., & Hager, J. C. Deliberate facial movements. *Child Development*, 1980, 51, 886-891.
- Izard, C. E. *The face of emotion*. New York: Appleton-Century-Crofts, 1971.
- Izard, C. E. Differential emotions theory and the facial feedback hypothesis of emotion activation: Comments on Tourangeau and Ellsworth's "The role of facial response in the experience of emotion." *Journal of Personality and Social Psychology*, 1981, 40, 350-354.
- Orne, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 1962, 17, 776-783.
- Tomkins, S. S. *Affect, imagery, consciousness. Vol. 1. The positive affects*. New York: Springer, 1962.
- Tomkins, S. S. The role of facial response in the experience of emotion: A reply to Tourangeau and Ellsworth. *Journal of Personality and Social Psychology*, 1981, 40, 355-357.
- Tourangeau, R., & Ellsworth, P. C. The role of facial response in the experience of emotion. *Journal of Personality and Social Psychology*, 1979, 37, 1519-1531.

Received April 17, 1980 ■