Journal of Personality and Social Psychology
1980, Vol. 38, No. 5, 811–824

Nonverbal Behavior and the Theory of Emotion: The Facial Feedback Hypothesis

Ross Buck
University of Connecticut

The *facial feedback hypothesis,* that skeletal muscle feedback from facial expressions plays a causal role in regulating emotional experience and behavior, is an important part of several comtemporary theories of emotion. A review of relevant research indicates that studies reporting support for this hypothesis have, without exception, used within-subjects designs and that therefore only a restricted version of the hypothesis has been tested. Also, the results of some of these studies must be questioned due to demand characteristics and other problems. It is suggested that visceral feedback may make a more direct contribution to emotional processes than facial feedback does and that the "readout" functions of facial expressions are more important than any feedback functions.

Peripheral theories of motivation assume that important aspects of motivational states depend on peripheral signals and cues: dryness of mouth for thirst, stomach contractions for hunger, and so forth. These approaches have been challenged by research demonstrating that peripheral cues are neither necessary nor sufficient for the occurrence of motivated behavior. At most, such cues appear to function to help inform the central nervous system of the state of the organism. More recent theories have stressed the role of the central nervous system in directly monitoring bodily needs and affecting appropriate behavior (Cofer & Appley, 1964). At present, there is relatively little evidence for the importance of peripheral processes in directing and controlling motivational phenomena (Buck, 1976, p. 59).

However, the peripheral position in the theory of emotion has recently enjoyed an upswing of interest, stimulated in large part by research in nonverbal behavior. The center of the argument goes back to the controversy over the James–Lange theory of emotion. In 1884, William James (1884/1968) argued that peripheral bodily changes are essential to add an emotional quality to the perception of an event:

> Bodily changes follow directly the *perception* of the exciting fact, and . . . our feeling of these same changes as they occur *is* the emotion. . . . Without the bodily states following on the perception, the latter would be purely cognitive in form, pale, colorless, destitute of emotional warmth. (p. 19; italics in the original)

James (1884/1968) clearly felt that skeletal muscle activity, including facial expression, was involved in these changes:

> Can one fancy the state of rage and picture no ebullition of it in the chest, no flushing of the face, no dilation of the nostrils, no clenching of the teeth, no impulse to vigorous action, but in their stead limp muscles, calm breathing, and a placid face? (p. 23)

However, early discussions and criticisms of the theory focused on the role of visceral sensations mediated by the autonomic nervous system. Cannon (1932) based his famous critique of the James–Lange theory in part on the notions that (a) visceral sensation is too diffuse and insensitive to account for the wide range of human emotional experience and (b) the autonomic nervous system acts too slowly

The author thanks Reuben Baron and Robert E. Miller for their comments and suggestions on this article.

Requests for reprints should be sent to Ross Buck, Dept. of Communication Sciences, U-85, University of Connecticut, Storrs, Connecticut 06268.

to account for the speed and lability of human emotional experience.

In recent years, two kinds of major attempts have been made to answer these criticisms. One, suggested over 50 years ago by Bertrand Russell (1961), is that cognitions arising from the responder's understanding of the situation producing the bodily sensations account for the quality and lability of emotional experience. This notion is at the heart of Schachter's (1964) self-attribution theory of emotion, which holds that emotion is a joint function of autonomic arousal and cognitive attributions or "labels" for that arousal. From this point of view, Schachter was able to integrate much of the research on the James–Lange theory of emotion (cf. Buck, 1976). More recently, it has been pointed out that proprioceptive cues associated with skeletal muscle activity may be involved in addition to automatic arousal in the bodily sensations associated with emotion. This has led to extensions of Schachter's position in the general context of self-attribution theory (Bem, 1967, 1972; Heider, 1958; Kelly, 1967). The responder may use the information from his own facial expressions and other expressive behaviors, as well as visceral cues, to interpret and label his or her emotional experience. For example, Laird (1974) has suggested that visceral arousal leads to self-attributions of the intensity of emotion, whereas feedback from one's expressive behavior leads to self-attributions of the quality of emotion.

The other answer to the criticisms of the James–Lange theory is based on Darwin's contention, expressed in *The Expression of the Emotions in Man and Animals,* that emotional processes are directly and intimately related to emotional expression.

> Most of our emotions are so closely connected with their expression, that they hardly exist if the body remains passive. . . . The free expression of outward signs of an emotion intensifies it. On the other hand, the repression, as far as this is possible, of all outward signs, softens our emotions. (Darwin, 1872, quoted in Plutchik, 1962, p. 153)

From this has developed the notion that the skeletal muscle activity associated with emotion, and particularly facial expression, plays a direct role in regulating emotional processes. Ekman, Friesen, and Ellsworth (1972) argue that these muscles are sufficiently differentiated and respond sufficiently quickly to underlie emotional experience: "The face might . . . fill the informational gap left by a solely visceral theory of emotion, distinguishing one emotion from another, changing rapidly and providing feedback about what is occurring to the person" (p. 173). This neo-Darwinian position has been adopted by a number of recent theories of emotion that are in other ways quite different from one another, including the theories of Plutchik (1962), Tomkins (1962, 1963), Gellhorn (1964), and Izard (1971, 1977).

The Facial Feedback Hypothesis

Collectively, these theories suggest what might be termed the *facial feedback hypothesis*: that facial expressions provide feedback to the responder that is necessary or sufficient to affect significantly his or her emotional experience and behavior. However, none of these theories has dealt with this specific hypothesis in sufficient detail. This article will critically examine the facial feedback hypothesis and suggest that two versions of the hypothesis—a between-subjects version and a within-subjects version—can be constructed from the existing discussions. There is no evidence for the former version and considerable evidence against it, whereas the several studies appearing to support the latter version are flawed in certain respects and are unconvincing. We shall suggest that there is sufficient evidence of the importance of *autonomic* feedback to support a cognitive–interpretive self-attribution theory such as that of Schachter's, but that there is as yet no convincing empirical evidence that facial feedback contributes to such a process. Instead, we shall argue that facial expressions serve a "readout" function as a channel of social communication and that their feedback function is at best of limited importance.

Two Versions of the Hypothesis

The facial feedback hypothesis occupies a prominent place in recent theories of emotion. However, it has never been stated clearly and

its implications have never been fully explored. The heart of the facial feedback hypothesis is the causal assertion that feedback from facial expressions affects emotional experience and behavior. One of the central implications of this would seem to be that the degree of feedback should be positively related to other indices of emotion—that the greater the feedback, the greater the emotional response. Since virtually by definition facial feedback must be related to facial expression, such expression should likewise be positively related to other emotional indices. These implications are clearly stated in the above quotation from Darwin that emotions "hardly exist if the body remains passive" and in Lanzetta, Cartwright-Smith, and Kleck's (1976) statement that "awareness of one's facial expressions *is* the emotion" (p. 371).

Given that the facial feedback hypothesis implies a "positive relationship" between facial expression and other indices of emotion, one must then inquire about how that relationship is assessed. Is it an intersubject relationship, that is, do persons who show much expression have a greater emotional response on other indices than other persons who show little expression? This seems to be implied in Darwin's statement. Or is it an intrasubject relationship, that is, will a given individual show a greater emotional response on an occasion when he or she is expressive than when he or she is not? Or do both relationships apply?

Unfortunately, the existing discussions of the facial feedback hypothesis do not make the answers to these questions clear, and one may in fact construct from them two versions of the hypothesis. One, which we shall descriptively term the "between-subjects version" of the hypothesis, holds that individual differences in the feedback from emotional expression are related to individual differences on other emotional indices: that for a given emotional stimulus, Person A who freely expresses his or her feelings will have a greater response on other indices of emotion than will a different Person B who shows little expression. The other "within-subjects version" of the facial feedback hypothesis states that Person A confronted by a given emotional stimulus will have a greater response on other affective indices if he or she freely expresses an emotion than if *that same Person A* were to show little expression. These two versions of the facial feedback hypothesis appear similar because both involve the relationship between facial expression and other indices of affective response. However, the between-subjects version is a statement of an intersubject relationship, whereas the within-subjects version is a statement of an intrasubject relationship. Statistically and logically, intersubject correlation does not predict intrasubject correlation and vice versa.

This is illustrated in Tables 1 and 2, which show hypothetical facial and physiological responses of Persons A and B to two affective stimuli. They also illustrate that if one makes the common assumption that stronger emotional stimuli lead to larger responses on all indices of affect, the between-subjects version implies the within-subjects version, whereas the opposite is not the case. Table 1 shows the between-subjects version of the hypothesis: Person A has a larger facial reaction than Person B, regardless of the nature of the affective stimulus, and therefore Person A's physiological responses are also larger. For the stronger emotional stimulus (Stimulus 2), both A and B show larger responses than they do with Stimulus 1. The result is that both intersubject and intrasubject relationships between the facial and physiological responses are positive.

Table 2 shows a situation where the within-subjects version of the facial feedback hypothesis holds, but the between-subjects does not. Person A is more facially expressive but has a smaller physiological response than Person B, resulting in a *negative* intersubject relationship between facial expression and physiological response. However, both A and B have larger facial reactions and consequently larger physiological responses to Stimulus 2 than they do to Stimulus 1, resulting in a positive intrasubject relationship. Thus, the within-subjects version of the facial feedback hypothesis does not imply the between-subjects version.

The following paragraphs will review the evidence relative to the between-subjects and within-subjects versions of the facial feedback hypothesis.

Table 1
Between-Subjects Version of the Facial Feedback Hypothesis Showing a Positive Intrasubject Relationship Between Facial and Physiological Response

Individual	Stimulus 1	Stimulus 2	Average showing intersubject relationship
Person A			
Facial activity	Moderate	High	High–moderate
Physiological *R*	Moderate	High	High–moderate
Person B			
Facial activity	Low	Moderate	Low–moderate
Physiological *R*	Low	Moderate	Low–moderate
Average showing intrasubject relationship			
Facial activity	Low–moderate	High–moderate	—
Physiological *R*	Low–moderate	High–moderate	—

The Between-Subjects Version

The between-subjects version of the facial feedback hypothesis implies that a lack of overt expression somehow short-circuits the whole emotional process. A nonexpressive person should be lower than an expressive person on all indices of emotion—behavioral, self-report, and physiological—when both are confronted by a given affective stimulus. To assess the validity of this hypothesis, one would investigate differences in the emotional response of individuals who differ in facial feedback. The most obvious and convenient, if not the most powerful, study of this kind for use with humans would be correlational in nature: Affective stimuli would be presented to expressive and nonexpressive persons and their reactions would be assessed, with the expectation that nonexpressive persons would show fewer indications of emotion on all indices. Experimental procedures could be approximated in humans by studying persons with different degrees of facial paralysis, with the expectation that measures of emotional responding would vary inversely with the degree of paralysis. Experimental techniques could also be applied in animals, particularly primates; the hypothesis would imply that drug or lesion-induced facial paralysis should interfere with the acquisition, if not the retention, of emotional behavior.

Studies analogous to these have appeared during the controversy over the James–Lange theory, but they have been directed toward the evaluation of the affective contribution of visceral rather than facial feedback. For example, Hohmann (Note 1) studied the reported emotional experiences of patients with spinal cord injuries and demonstrated that patients with higher spinal lesions (and thus a greater loss of bodily sensation) reported a more cognitive kind of emotional experience. ("It's a mental kind of anger." "It doesn't have the heat to it that it used to.") Similarly, Delgado (1969) observed that a patient who had undergone a unilateral sympathectomy reported that he could no longer be thrilled by music on the sympathectomized side of his body, whereas his response on the other side was unchanged. Also, animal studies have shown that bilateral sympathectomy retards the acquisition of emotional behavior, although it has no effects on the retention of learned emotional responses (Wynne & Solomon, 1955). These studies have demonstrated convincingly that visceral feedback normally plays an important role in emotional processes, although it is apparently neither necessary nor sufficient for *all* kinds of emotional experience or behavior (Buck, 1976, pp. 42–49).

Similar kinds of studies could evaluate the validity of the between-subjects version of the facial feedback hypothesis in explaining emotional experience and behavior. However, none of the recent studies purporting to support the facial feedback hypothesis has used the between-subjects paradigm necessary to evaluate the between-subjects form of the hypothesis. In fact, the available evidence bearing on the between-subjects form of the facial feed-

Table 2
Within-Subjects Version of the Facial Feedback Hypothesis Showing a Negative Intersubject Relationship Between Facial and Physiological Responses

Individual	Stimulus 1	Stimulus 2	Average showing intersubject relationship
Person A			
Facial activity	Moderate	High	High–moderate
Physiological *R*	Low	Moderate	Low–moderate
Person B			
Facial activity	Low	Moderate	Low–moderate
Physiological *R*	Moderate	High	High–moderate
Average showing intrasubject relationship			
Facial activity	Low–moderate	High–moderate	—
Physiological *R*	Low–moderate	High–moderate	—

back hypothesis appears to be negative (cf. Buck, 1979). Correlational studies on the intersubject relationship between facial/gestural expressiveness and physiological responding indicate that facially nonexpressive persons tend to have *larger* skin conductance and heart rate responses to emotional stimuli than do expressive persons (Buck, 1977; Buck, Miller, & Caul, 1974; Buck, Savin, Miller, & Caul, 1969, 1972; Lanzetta & Kleck, 1970; Miller, 1974). Also, Waid (1976) has summarized a number of studies suggesting that skin conductance responding is positively related to the degree of socialization, and it seems unlikely that highly socialized persons are more overtly emotional than less socialized persons. In fact, one might expect the opposite to be the case.

The studies on the relationship between overt emotional expressiveness and physiological responding that have used between-subjects designs have tended to support H. E. Jones' (1950) distinction between "externalizing," where a person is outwardly expressive but has small skin conductance responses, versus "internalizing," where one is less outwardly expressive but has large skin conductance responses. Conceptually, the externalizing–internalizing phenomenon seems inconsistent with the facial feedback hypothesis insofar as the latter implies that increased expression should cause an increased response on other emotional indices. Izard (1977) has noted that facial feedback may occur without overt expression—he cites, for example, the Schwartz (Schwartz, Fair, Greenberg, Freedman, & Klerman, 1974; Schwartz, Fair, Salt, Mandel, & Klerman, 1976) demonstrations of "covert expression" showing that electromyograph recordings can reveal the occurrence of facial muscle movement not visible on the face. However, this does not alter the implication that the degree of expression should in general be positively related to other emotional indices.

The internalizing–externalizing distinction has been discussed in detail elsewhere (see Buck, 1976, pp. 259–274; 1979). The remainder of this article will discuss the within-subjects version of the facial feedback hypothesis.

The Within-Subjects Version

The within-subjects version of the facial feedback hypothesis states that a given individual will have a greater emotional response on other indices if he or she is expressive than if he or she is not. At first glance, this appears similar to the between-subjects version of the hypothesis, and it is all too easy to discuss the "relationship" between emotional expression and other affective indices without making clear whether one is considering the intersubject or intrasubject relationship. However, the two in fact are quite different: One does not predict the form of the other either mathematically or conceptually, they involve different assumptions about the mechanisms involved, they have different practical implica-

tions, and they require different kinds of research design and analysis to test their validity. Their apparent similarity has thus caused considerable confusion.[1]

The difference between the between-subjects and within-subjects versions of the facial feedback hypothesis was demonstrated in Table 2 by the fact that it is possible simultaneously to have negative intersubject and positive intrasubject relationships between emotional expression and other affective indices. Buck et al. (1974) in fact found this pattern of results: Both a negative intersubject correlation between expressiveness measures and skin conductance responding (average $r = -.50$, $z = 2.92$, $p < .005$), which is contrary to the between-subjects version of the facial feedback hypothesis, and a positive intrasubject correlation between rated expressiveness and skin conductance responding (average $r = .15$, $z = 3.75$, $p < .001$), which is consistent with the within-subjects version of the hypothesis. In other words, Person A apparently has a larger skin conductance response when he or she is expressive than when he or she is not, but Person A who *in general* is expressive tends to have a *smaller* skin conductance response than Person B who in general is not expressive.

A number of studies have appeared in recent years that have been cited as supporting the facial feedback hypothesis. None of these differentiates between the between-subjects and within-subjects versions of the hypothesis. Some investigators, in making the general statement that their studies involve the relationship between the expression of emotion and other indices of emotional state, imply that their findings should be applicable to the between-subjects version of the facial feedback hypothesis. However, these studies without exception used within-subjects designs, and their findings cannot be used to evaluate a hypothesis concerning a between-subjects phenomenon.

Once the confusion about the between-subjects and within-subjects versions of the facial feedback hypothesis has been cleared, it becomes possible to examine the explanatory power and usefulness of the remaining within-subjects version of the hypothesis that has been the subject of considerable experimentation. One can ask, for example, how important facial feedback normally is in the emotional life of the individual. The evidence bearing on the between-subjects version of the hypothesis shows that nonexpressive persons have strong skin conductance and heart rate reactions to affective stimuli, implying that facial feedback may not be necessary for emotion to occur. It remains to be demonstrated whether facial feedback is as effective as visceral feedback has been shown to be in affecting emotional experience. It is still possible that facial feedback may be sufficient to alter, or even create, emotional experience and behavior. The succeeding paragraphs will review the studies relevant to the within-subjects version of the facial feedback hypothesis in an attempt to address these issues.

Experimental Evidence

The first experiments relevant to the within-subjects version of the facial feedback hypothesis stemmed from self-attribution approaches. Bandler, Madaras, and Bem (1968) demonstrated that subjects who "observed themselves" escaping from shocks rated those

---

[1] It might be noted that the within-subjects version of the hypothesis may be stated in a strong version that holds across situations and a weak version that pertains to only one situation. The former states that Person A will have a greater response on other affect indices in a situation where A is overtly expressive than in a *different situation* where A is not expressive. For example, if a given male is overtly expressive in an aggressive situation but not in an emotional situation (i.e., viewing affect-laden color slides), one would expect that he would have greater physiological responses in the aggressive situation than he would in the emotional. The latter version states that *within a given kind of affective situation,* Person A will have a greater response on other affect indices when A is expressive than when A is not. For example, if a male is overtly expressive to one affect-loaded color slide but not another, he should also have a greater physiological response to the former slide than to the latter. There does not appear to be any evidence for or against the stronger "across situation" version of the within-subjects hypothesis, as no study has yet investigated overt expressiveness and other affect indices in the same person in qualitatively different situations: All have involved only one kind of affective situation. This article therefore will not consider it further (cf. Buck, 1979).

shocks as more uncomfortable than physically identical shocks that they endured without escaping. Kopel and Arkowitz (1974) found a similar result in an experiment in which subjects role played being upset or calm when given shocks. The subjects who role played an upset reaction showed a subsequent decrease in their pain and shock tolerance thresholds relative to their performance before role playing, whereas subjects who role played a calm reaction showed higher thresholds to the shocks (r_m = .296 and .515, respectively).[2] Physiological responses (skin conductance in Bandler et al., 1968; heart rate in Kopel & Arkowitz, 1974) were not significantly affected by the role playing manipulations in these experiments.

The first studies directed specifically at evaluating the facial feedback hypothesis were reported by Laird (1974), who (like the above investigators) couched his discussion in the terms of self-attribution theory. However, he also pointed out that the neo-Darwinian theories such as those of Tomkins, Izard, and Gellhorn would make similar predictions. Laird attempted to manipulate facial expressions without manipulating cognitive awareness of the expressions or their relevance to the experiment. He did this by telling subjects to contract certain facial muscles that would result in facial expressions without specifically asking them to make those expressions. Thus for a "frowning" face, subjects were asked to bring their eyebrows down and together and to contract the jaws; for a "smiling" expression they were to draw the corners of the mouth back and up. The subjects watched and rated their affective responses to slides designed to elicit mild positive and negative reactions (slides of children playing vs. slides of Ku Klux Klan members) while "frowning" or "smiling." As expected (Laird, 1974), subjects reported feeling more angry when "frowning" and more happy when "smiling," although the effect of the expressions was weaker than the effects of the two types of pictures: "The differences between the means of the two picture conditions were 2½ to 7 times as large as the differences between expression condition means, despite the fact that the pictures were not by any means the most emotionally evocative." (p. 481; no r_m could be calculated from the data reported.) In a second study, subjects rated cartoons they viewed while "smiling" as funnier than cartoons viewed while "frowning." These results were interpreted as supporting the contention that the self-observation of one's expressive behavior is one determinant of the quality of one's emotional experience.

Laird showed a particular sensitivity to possible experimenter bias and other demand characteristics in these experiments. In the first study, he provided a control group of subjects who received the instructions with the experimental subjects but did not make the actual expressions while viewing the slides. The ratings of these control subjects were not significantly affected by the instructions. Although Laird acknowledged that this procedure was not foolproof, it did provide some evidence that the results were not due to experimenter bias. Also, the experimental subjects in both of Laird's studies were given a formal postexperimental questionnaire to assure that they were not aware of the actual purpose of the manipulation of their faces. A considerable number of subjects were judged to be aware and were therefore excluded from the analysis, including 16% of the experimental subjects in the first study and 19% of the subjects in the second.[3]

---

[2] Effect size estimates, based on r_m, are given when possible (Cohen, 1977; Friedman, 1968; Rosenthal, 1976).

[3] More recently, Laird has been studying individual differences in subjects' response to the expression manipulation, expecting such differences to relate to styles of self-attribution. Duncan and Laird (1977) studied the relationship between subjects' response to the expression manipulation and their response to a forced-compliance manipulation. They found that responses on the two tasks were related and interpreted the data in terms of individual differences in self-attribution styles. Also, Laird and Crosby (1974) found intersubject relationships between the response to an expression manipulation and questionnaire items assumed to reflect the use of situational versus self-produced cues. These studies are not directly relevant to the evaluation of the facial feedback hypothesis: Essentially they imply that facial feedback may be used by some persons but not by others. However, both of these studies suffered from high attrition rates due to awareness of the purposes of the experiment—35% in the former study and 19% in the latter—so that any results must be viewed with caution.

Lanzetta, Kleck, and their colleagues have conducted a series of experiments on the facial feedback hypothesis that Ekman and Oster (1979) consider to have shown the "strongest evidence for a positive link between voluntary facial expression and emotional experience" (p. 546). Unlike the self-attribution theorists, Lanzetta, Cartwright-Smith, and Kleck (1976) hold the neo-Darwinian position that there is a "direct link between facial expressive behavior and the processes responsible for autonomic arousal and subjective experience" and imply that the "awareness of one's facial expression *is* the emotion" (p. 369; p. 371). Laird (1974) and Lanzetta et al. (1976) have discussed the relative merits of the direct versus self-attribution positions: This topic will not be treated here.

Like the Bandler et al. (1968) and Kopel and Arkowitz (1974) studies, the experiments of Lanzetta and Kleck's group have used electric shocks as affective stimuli. Lanzetta et al. (1976) discussed three separate experiments. In the first two, subjects were given a "baseline" series of shocks varying in intensity. Each shock was preceded by a signal that indicated its intensity. The skin conductance response to the signal and the subjects' self-reports of shock painfulness were assessed. In the first study, following the baseline shocks, subjects were asked to "hide" any overt response to the signal, and another series of shocks was administered. In the second study, subjects were instructed to "hide" or "reveal" their overt response to the signal in a complex, although not counterbalanced, sequence. The skin conductance and self-report responses to the anticipation and reception of shocks under these instruction conditions were compared to the previous response to the baseline series. Unfortunately, the results of these two studies are impossible to interpret because of the noncounterbalanced nature of their designs. In both cases, responding during a baseline phase was compared to responding during a subsequent manipulation phase, so that habituation and other possible time-series effects were confounded with the effects of the manipulations.

In the third study, the instructions to the subjects were presented in a random order, allowing comparisons between the instruction conditions. The subjects were told to pose anticipating and receiving one of two levels of shock, either no shock at all or "extremely intense, almost unbearable shock." Actually, three levels of shock were delivered, defined as 33%, 66%, and 99% of the subject's previously established tolerance limit. A color slide signaled the subject regarding both the level of shock and the posing condition that he or she was to use on a given trial: A red 1 indicated that a low shock was to be given but that an intense shock should be posed, a green 2 indicated a moderate shock and a pose of no shock, and so forth.

The results indicated that when the subjects posed intense shocks, they rated the shocks as more intense ($r_m = .769$) and had larger skin conductance responses both to the signals and to the shocks themselves ($r_m = .716$ and .866, respectively). Lanzetta et al. concluded that the results were strongly supportive of the facial feedback hypothesis. Subsequent experiments by Kleck et al. (1976) using male subjects further supported these conclusions, showing a decrease in skin conductance responding ($r_m = .462$ in Study 1, .429 in Study 2) and self-reported pain ($r_m = .506$ in Study 1, .726 in Study 2) when the attenuation of expressive behavior was caused by social cues —the instruction that the subject was being observed while receiving shocks—rather than by direct instructions to modify expressive behavior.

In a recent experiment, Colby, Lanzetta, and Kleck (Note 2) had male subjects endure a pain tolerance test in which shock intensity was increased to the subjects' tolerance limit, whereupon he pushed a button to escape. Subjects were asked to pose being in high, moderate, or no pain during these trials, with the expectation that a lack of overt expression would be associated with lower skin conductance responding and a higher pain tolerance. Only the first of these hypotheses was confirmed: Skin conductance was highest when the subjects posed high pain ($r_m = .548$), but pain tolerance scores were not significantly affected by the posing conditions.

Detailed reports of other studies relevant to the within-subjects version of the facial feedback hypothesis have not yet appeared in the experimental literature. For example, Ar-

gyle (1975) reports Shimoda's finding that if a person adopts a facial expression during interaction, his or her mood tends to change accordingly, and Izard, Kidd, and Kotsch (Note 3) have reported that subjects instructed to hold their faces in positions characteristic of given emotions tended to rate their subjective experience accordingly. Also, Kleck and Lanzetta (Note 4) reported that Lanzetta and Johnson altered the mood ratings of subjects via a "biofeedback" procedure that induced them to smile or frown. On the other hand, Ekman and Oster (1979, p. 546) report that Tourangeau and Ellsworth, using facial manipulations, failed to show significant effects on self-reported emotion and found ambiguous effects on physiological responses. Although this result is not directly related to emotion, it might be noted that Cartwright-Smith (Note 5) has recently made the interesting finding that an exaggerated facial display of effort (grimace) may increase the strength of the grip in the forearm.

Evaluation

In evaluating whether these studies offer convincing support for the within-subjects version of the facial feedback hypothesis, several problems stand out. First, all of these studies used highly artificial experimental situations and employed deception to some extent. One of the major dependent variables of these experiments—self-reported emotional experience—is particularly susceptible to experimenter bias and other demand characteristics that are difficult to control in such studies. Only the experiment by Laird (1974) demonstrated an open concern with this possibility, and the measures he took to guard against it led to a high rate of subject loss that was not found in the other experiments. One must question, therefore, the degree to which demand characteristics may have affected the self-reports in these studies. In adopting a given expression, a subject might consciously or unconsciously conclude that a corresponding emotion is desired by the experimenter and might respond accordingly. This explanation cannot, however, account for the physiological findings.

Another problem with these studies is that although they all generally supported the within-subjects version of the facial feedback hypothesis, their results were not entirely consistent with one another. For example, Lanzetta and Kleck's group has consistently reported that expression manipulations affect physiological responding, whereas the studies of Bandler et al. (1968) and Kopel and Arkowitz (1974) did not. Also, the finding by Colby et al. (Note 2) that an expression manipulation did not influence pain tolerance is not consistent with the studies showing that pain and shock tolerance thresholds and reports of shock painfulness are affected by such manipulations. The explanations for these inconsistencies are not clear.

A third problem concerns the studies employing experimenter-induced facial movement, such as those of Laird and his colleagues. Izard (Note 6) has had difficulty using this technique with emotions other than happiness and anger, and he reports that Kotsch found an increase in anger ratings using non-emotion-related facial contractions. Izard suggests from his experience that only when "subjects initiate the facial patterns for reasons of their own, the reasons consonant with the desired effect," will the facial muscle movements play a role in emotion regulation. A related potential problem with experimenter-induced feedback, which is acknowledged by Laird (1974), is that such expressions must be held for an unnaturally long period of time. The feedback from such unnatural expressions could be discounted by the subject, thus working against the hypothesis (Ekman & Oster, 1979). Alternately, such extended feedback might be stronger and more obvious to the subject than that from natural expression, thus leading to an inaccurate confirmation of the hypothesis.

A final problem with these studies is that all of those employing "subject generated movements" have used electric shock as an affective stimulus. Some would argue that the affective state induced by shock is not clearly emotional, and in any case such a reliance on a single kind of eliciting stimulus is hazardous.

All in all, it could be argued that these studies do not convincingly demonstrate that facial feedback is a sufficiently important phe-

nomenon in the emotional life of the individual to "fill the information gap left by a solely visceral theory of emotion." Nevertheless, they do suggest that the within-subjects version of the hypothesis is valid under some conditions. The evidence that facial expression can apparently affect skin conductance responding under some conditions is particularly interesting. Kleck et al. (1976) suggest that this may be due to the conditioning experiences of the individual over his or her lifetime. They argue that both expressive behavior and autonomic nervous system responses are unconditioned responses to affective stimuli and that the former typically precede the slow-acting physiological responses. The internal cues associated with the expressions could come to function as conditioned stimuli for the autonomic nervous system responses, so that they could come to evoke those responses in the absence of the affective unconditioned stimulus.

However, there is another factor that must be considered by any theory that regards facial feedback as playing a causal role in emotional phenomena. This, as Ekman and Friesen (1969, 1975) have shown, is that much facial behavior is not associated with external affective stimuli but is instead related to social and situational demands. Our long experience of "lying with our faces" in conformance with such demands must lead us to disregard our own facial feedback as a reliable source of information to some extent. In the terminology of conditioning theory, our facial expressions must occur so often in the absence of the unconditioned stimulus that a considerable amount of extinction must take place. Ekman and Friesen's (1969) finding that facial movements are more controlled than other body movements would suggest that such extinction would be particularly common with facial expressions. However, the organism might learn to discriminate between the internal cues associated with spontaneous facial expressions, which tend to be reliably linked to actual affective stimuli, and those associated with intentional facial expressions, which do not. This may account for the fact, according to Kleck and Lanzetta (Note 4), that posed facial expression in the absence of any affective stimulus fails to cause physiological arousal.

Implications for the Theory of Emotion

The results of the studies on the facial feedback hypothesis do not appear to compel the acceptance of a peripheral theory of emotion: There is nothing to suggest that the central nervous system is not responsible for the overall control of the affective process. However, it is clear from the studies of the effects of visceral feedback that the affective process is normally accompanied and enriched by peripheral feedback. The veteran with a spinal cord injury and the patient with a unilateral sympathectomy report that their experience of emotion is incomplete. Peripheral factors, including facial and other bodily feedback, normally contribute to the emotional state, and under the right circumstances they may alter the emotional state.

Peripheral Feedback and Conditioning

It may well be that proprioceptive and interoceptive conditioning are central to the process by which peripheral events become able to alter affective states. We carry around within us a large variety of internal visceral and proprioceptive stimuli associated with digestion, circulation, posture, facial expression, sexual functioning, and so forth, that can become conditioned stimuli and responses associated both with each other and with external stimuli. For example, the internal stimuli associated with eating and digestion can become conditioned stimuli producing emotional responses similar to those generally produced by the eating situation. Thus a child might consistently receive affection while eating, which is an unconditioned stimulus for comfortable affect. Since the internal stimuli related to digestion are consistently associated with affection, these internal stimuli may become conditioned stimuli producing comfortable affect, and the individual may come to derive unusual comfort from eating. Similarly, if a child's parents consistently fight at the table, the same kinds of internal stimuli could become conditioned stimuli for fearful and angry emotions. An interesting and important aspect of this analysis is that the latter child could not escape from these internal stimuli in the same way that he or she could escape from other external stimuli that became con-

ditioned to cause negative affects (cf. Buck, 1976, pp. 115–122).

Razran (1961) points out that these internal stimuli, based as they are on recurring bodily functions, are always available for conditioning and are ever present afterward to reflect the results of that conditioning. Moreover, due to the relative paucity of sensory fibers from the internal organs, they are generally not consciously reportable. As a result, a complex, unique, and largely unconscious system of conditioned internal reactions must be built up within the life of each individual. One might expect that "facial feedback" would participate in this conditioning process, with the stipulation that intentional facial displays may be less reliably linked to other stimuli than are spontaneous displays.

One of the interesting features of this conditioning approach is that it has been advanced to explain both the positive correlation between facial expression and skin conductance responding found in studies using within-subjects designs and the negative relationship between these variables found in studies using between-subjects designs. As noted above, Kleck et al. (1976) suggested that expressive responses may come to function as conditioned stimuli for autonomic responses because of their common associations with the unconditioned affective stimuli and the time difference between them. The result would be a positive relationship between expression and autonomic responding when considered within the conditioning history of a single individual. Similarly, Buck et al. (1974) invoked a conditioning explanation for the negative relationship between facial expressiveness and skin conductance responding. It was reasoned that the social learning experiences associated with inhibiting an overt affective response may tend to involve stressful social rebukes that are unconditioned stimuli eliciting autonomic arousal. The inhibitory responses may become conditioned stimuli eliciting similar autonomic responses because of their association with these experiences. If a given individual endures many of these unpleasant social learning experiences, he or she will come to mask facial expression but will display large autonomic responding in affective situations. Another person who endures fewer inhibitory experiences will have more facial expression and less autonomic responding. The result would be a negative relationship between expressiveness and autonomic responding between subjects. Thus a consideration of proprioceptive and interoceptive conditioning may suggest the conditions under which facial and other kinds of peripheral feedback may come to alter other indices of affect.

Facial Behavior as a Readout Device

The conclusion that facial feedback may play a secondary role in affective processes does not imply that facial expression is an epiphenomenon with little importance to the study of emotion. It could be argued that the real importance of facial expressions and other nonverbal behaviors lies not in their role as a feedback device but in their role as a readout device. Our facial expressions reflect central processes, not the reverse. That, some ethologists have argued, is what they are there for. For example, Andrew (1965) has suggested that the use of facial expression and other kinds of communicative displays is a function of the extent of socialization of the given species and the communication demands of the environment, with the consequent need to coordinate social behavior. Social species show a much greater range of facial expressions than do species that lead solitary lives, and species that live in settings requiring much communication (e.g., the plains-dwelling baboon) show more sophisticated repertories of nonverbal displays than do similar species with other lifestyles (e.g., the forest-dwelling mandrill and drill baboon).

Due to the importance of the affective readout in the coordination of social behavior, affective responses that differ in their "visibility"—the degree to which they are normally apparent to the responder and to observers—must undergo both different evolutionary histories within the species and different social learning histories within the individual. Responses that are relatively "invisible" to others, including most autonomic responses, should be relatively unaffected by selection pressures favoring an affective readout, while more visible responses may be magnified in the service of social coordination. Since Dar-

win (1872), it has been suggested that affective displays evolve by the elaboration of some visible response that tends to occur naturally in a given affective situation. This process of "ritualization" has been described by a number of writers (Blest, 1961; Tinbergen, 1959). For example, the use of the eyebrow flash in human displays has been held to be the result of the ritualization of the eyebrow lift associated with the widening of the eyes that naturally accompanies an increase in visual attention (Eibl-Eibesfelt, 1972).

Affective responses that differ in visibility must also undergo different conditioning histories within the life of an individual (Buck, 1971). External agents of socialization are able to subject visible displays, including facial expressions and other nonverbal behaviors, to rewards and punishments that may cause them to be intensified or masked. Also, an individual has the opportunity to observe the visible responses of others to serve as models for his or her own behavior. As a result, visible affective responses come under the influence of a variety of "display rules" (Ekman & Friesen, 1975). This is not true with the less visible autonomic responses. The autonomic response in a given affective situation is normally a consequence of the existence of stimuli that evoke conditioned or unconditioned autonomic responses. The responder and others in the situation are not ordinarily aware of such responses and cannot control them unless special equipment (i.e., biofeedback devices) is employed.

Because of the difference in the visibility of expressive nonverbal behavior and autonomic responses and the consequent difference in their social learning histories, it is perhaps not surprising that they seem to be related to one another in complex ways. Also, it can be seen that facial expressions and other nonverbal behaviors not only give others a more informative readout of the affective state of the individual but also allow that individual greater control over such a readout. We try not to show our happiness, surprise, or anger when it is not appropriate to do so. Such a *controlled* display of affect is, and perhaps has long been, necessary for the smooth functioning of primate social organizations. Such displays may have also resulted in more complex facial feedback to the responder, but this seems to have been secondary to their readout functions.

Conclusions

This article argues that the evidence for theories positing that facial feedback has a major causal role in emotional processing is unconvincing. Considerable confusion has resulted from the failure of these theories to examine the implications of the facial feedback hypothesis fully, particularly the difference between what are here termed the between-subjects and the within-subjects versions of the hypothesis. There is evidence that in within-subjects designs, posing different emotional expressions leads to corresponding changes on other affective indices, but many of the findings, particularly those involving subjective report, may be due to demand characteristics. At present there is insufficient evidence to conclude that facial feedback is either necessary or sufficient for the occurrence of emotion, and the evidence for *any* contribution of facial feedback to emotional experience is less convincing than the evidence for visceral feedback. It is argued instead that facial expression has evolved in humans as a means of affective communication and that facial expressions and other nonverbal behaviors provide a controlled "readout" of central affective processes. The basic central affective processes have perhaps not changed greatly through mammalian evolution, but the ability to express them has.

Reference Notes

1. Hohmann, G. W. *The effect of dysfunctions of the autonomic nervous system on experienced feelings and emotions.* Paper presented at the Conference on Emotions and Feelings, New School of Social Research, New York, 1962.
2. Colby, C. A., Lanzetta, J. T., & Kleck, R. E. *Effects of the expression of pain on autonomic and pain tolerance responses to subject-controlled shock.* Paper presented at the meeting of the Eastern Psychological Association, Boston, April 1977.
3. Izard, C. E., Kidd, R. F., & Kotsch, W. E. *The facial expression of emotion as an activator of the subjective experience of emotion.* Unpublished manuscript, Vanderbilt University, 1974.

4. Kleck, R., & Lanzetta, J. *The role of expressive behavior in emotional processes.* Paper presented at the National Conference on Body Language, New York, 1977.
5. Cartwright-Smith, J. *The grimace: Facilitating forearm strength by exaggerating the nonverbal facial display of effort.* Paper presented at the meeting of the Eastern Psychological Association, Philadelphia, Pa., April 1979.
6. C. E. Izard, Personal communication, December 12, 1977.

References

Andrew, R. J. The origins of facial expressions. *Scientific American,* 1965, *213*(4), 88–94.

Argyle, M. *Bodily communication.* New York: International Universities Press, 1975.

Bandler, R., Madaras, G. R., & Bem, D. J. Self-observation as a source of pain perception. *Journal of Personality and Social Psychology,* 1968, *9,* 205–209.

Bem, D. J. Self perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review,* 1967, *74,* 183–200.

Bem, D. J. Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6). New York: Academic Press, 1972.

Blest, A. D. The concept of ritualization. In W. H. Thorpe (Ed.), *Current problems in animal behaviour.* Cambridge, England: Cambridge University Press, 1961, 102–124.

Buck, R. Differences in social learning underlying overt-behavioral, self-report, and physiological responses to emotion. *Research in Education,* 1971, *6,* 17. (Abstract)

Buck, R. *Human motivation and emotion.* New York: Wiley, 1976.

Buck, R. Nonverbal communication accuracy in preschool children: Relationships with personality and skin conductance. *Journal of Personality and Social Psychology,* 1977, *33,* 225–236.

Buck, R. Individual differences in nonverbal sending accuracy and electrodermal responding: The externalizing–internalizing dimension. In R. Rosenthal (Ed.), *Skill in nonverbal communication: Individual differences.* Cambridge, Mass.: Oelgeschlager, Gunn & Hain, 1979.

Buck, R., Miller, R. E., & Caul, W. F. Sex, personality, and physiological variables in the communication of affect via facial expression. *Journal of Personality and Social Psychology,* 1974, *30,* 587–596.

Buck, R., Savin, V. J., Miller, R. E., & Caul, W. F. Nonverbal communication of affect in humans. *Proceedings of the 77th Annual Convention of the American Psychological Association,* 1969, *4,* 367–368. (Summary)

Buck, R., Savin, V. J., Miller, R. E., & Caul, W. F. Communication of affect through facial expression in humans. *Journal of Personality and Social Psychology,* 1972, *23,* 362–371.

Cannon, W. B. *The wisdom of the body.* New York: Norton, 1932.

Cofer, C. N., & Appley, M. H. *Motivation: Theory and research.* New York: Wiley, 1964.

Cohen, J. *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press, 1977.

Darwin, C. *The expression of emotions in man and animals.* London: John Murray, 1872.

Delgado, J. M. R. *Physical control of the mind.* New York: Harper & Row, 1969.

Duncan, J., & Laird, J. Cross-modality consistencies in individual differences in self-attribution. *Journal of Personality,* 1977, *45,* 191–206.

Eibl-Eibesfeldt, I. Similarities and differences between cultures in expressive movements. In R. A. Hinde (Ed.), *Nonverbal communication.* Cambridge, England: Cambridge University Press, 1972.

Ekman, P., & Friesen, W. V. Nonverbal leakage and clues to deception. *Psychiatry,* 1969, *32,* 88–105.

Ekman, P., & Friesen, W. V. *Unmasking the face.* Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Ekman, P., Friesen, W. V., & Ellsworth, P. *Emotion in the human face.* Pergamon Press, New York. 1972.

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

Friedman, H. Magnitude of experimental effect and a table for its rapid estimation. *Psychological Bulletin,* 1968, *70,* 245–251.

Gellhorn, E. Motion and emotion: The role of proprioception in the physiology and pathology of emotions. *Psychological Review,* 1964, *71,* 457–472.

Heider, F. *The psychology of interpersonal relations.* New York: Wiley, 1958.

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

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

James, W. What is an emotion? In M. Arnold (Ed.), *The nature of emotion.* Baltimore, Md.: Penguin, 1968. (Originally published, 1884.)

Jones, H. E. The study of patterns of emotional expression. In M. Reymert (Ed.), *Feelings and emotions.* New York: McGraw-Hill, 1950.

Kelley, H. H. Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15). Lincoln: University of Nebraska Press, 1967.

Kleck, R. E., et al. Effects of being observed on expressive, subjective, and physiological reactions to painful stimuli. *Journal of Personality and Social Psychology,* 1976, *34,* 1211–1218.

Kopel, S. A., & Arkowitz, H. S. Role playing as a source of self-observation and behavior change. *Journal of Personality and Social Psychology,* 1974, *29,* 677–686.

Laird, J. D. Self-attribution of emotion: The effects of expressive behavior on the quality of emotional experience. *Journal of Personality and Social Psychology,* 1974, *29,* 475–486.

Laird, J., & Crosby, M. Individual differences in self-attribution of emotion. In H. London & R. Nisbett (Eds.), *Cognitive alteration of feeling states*. Chicago: Aldine-Atherton, 1974.

Lanzetta, J. T., Cartwright-Smith, J., & Kleck, R. E. Effects of nonverbal dissimulation on emotional experience and autonomic arousal. *Journal of Personality and Social Psychology*, 1976, *33*, 354–370.

Lanzetta, J. T., & Kleck, R. E. Encoding and decoding of nonverbal affect in humans. *Journal of Personality and Social Psychology*, 1970, *16*, 12–19.

Miller, R. E. Social and pharmacological influences on nonverbal communication in monkeys and man. In L. Krames, T. Alloway, & P. Pliner (Eds.), *Nonverbal communication*. New York: Plenum Press, 1974.

Plutchik, R. *The emotions: Facts, theories and a new model*. New York: Random House, 1962.

Razran, G. The observable unconscious and the inferable conscious in current Soviet psychophysiology. *Psychological Review*, 1961, *68*, 81–147.

Rosenthal, R. *Experimenter effects in behavioral research* (Enlarged ed.). New York: Irvington, 1976.

Russell, B. *An outline of philosophy*. Cleveland, Ohio: World, 1961.

Schachter, S. The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1). New York: Academic Press, 1964.

Schwartz, G. E., Fair, P. L., Greenberg, P. S., Freedman, M., & Klerman, J. L. Facial electromyography in the assessment of emotion. *Psychophysiology*, 1974, *11*, 237.

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

Tinbergen, N. Comparative studies of the behavior of gulls (Laridae): A progress report. *Behavior*, 1959, *15*, 1–70.

Tomkins, S. *Affect, imagery, and consciousness: Vol. 1. The positive affects*. New York: Springer, 1962.

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

Waid, W. N. Skin conductance response to both signaled and unsignaled noxious stimulation predicts level of socialization. *Journal of Personality and Social Psychology*, 1976, *34*, 923–929.

Wynne, L. C., & Solomon, R. L. Traumatic avoidance learning: Acquisition and extinction in dogs deprived of normal autonomic function. *Genetic Psychology Monographs*, 1955, *52*, 241–284.

Received June 4, 1979 ■