Basic and Applied Studies of
Spontaneous Expression Using the
Facial Action Coding System (FACS)

Edited by
PAUL EKMAN &
ERIKA L. ROSENBERG

Conclusion

What We Have Learned by Measuring Facial Behavior

PAUL EKMAN

This extraordinary collection of articles vividly illustrates how rich and diverse is the information that can be obtained by measuring facial behavior. Because measurement of facial behavior provides information about so many different phenomena, these chapters have appeared in many journals: 13 different journals are represented. It would be unusual for anyone to know all of these articles. I suspect that even the authors in this book are not familiar with all of the other articles published here. It was important, therefore, to collect in one place examples of the very diverse kinds of research that can benefit from measuring the face.

One might question why a personality psychologist needs to read about infant facial expression or a pain researcher should read about facial expression in psychotherapy. There are two answers. First, the face itself is a legitimate focus of study, and if one wants to learn what information can be derived from this multisignal system, then it is important to determine all the different kinds of information—messages—these signals can provide. It is not just students of the face who need to know the diverse literature represented here, but also students of emotion. It is no accident that each of these studies that measured facial behavior provides information relevant to students of emotion, for the face is a major source of information about emotion.

Unlike the research reported in this book, most of the psychological literature on facial expression did not measure the face but instead relied on observers' inferential judgments about what emotion is shown in facial behavior. Many of those studies focused on still photographs, most often posed still photographs. A reader of the psychological literature might well think that is all that has been done. But this book shows how much can be learned by actually measuring the face itself, not the inferences observers make when they see the face and by examining spontaneous facial behavior.

It is much more time-consuming to measure facial behavior itself—with FACS or any other technique (such as EMG, or Izard's MAX (1979)—than to obtain inferences about emotion from observers. However, none of the questions asked by these investigators could have been answered except by measuring the facial behavior. In my discussion of their work, I will point out why this is so using specific examples from this book.

I have organized by discussion by considering groups of articles, following the organization of the book. First, I will consider the substantive matters raised by the basic research chapters, and then the substantive issues raised by the applied research chapters. Then I will discuss methodological issues raised by various chapters.

Basic Research Issues

A number of very interesting theoretical issues are raised by the studies presented in this section. Three of the chapters considered individual differences in the magnitude of facial responses. Ekman, Friesen, and Simons wrote about hyper-startlers in response to a loud noise, and Ruch in one of his articles described hypo- and hyper-expressers in relation to humor, and in his other article reported individual differences in enjoyment smiles in relation to personality and mood measures. It is important to determine whether there are hypo- and hyper-expressers for each of the emotions, if such individual differences in the magnitude of facial responses are emotion specific or generalize across emotions, and whether such individual differences in the magnitude of expression are manifest also in the magnitude of subjective responses and physiological activity.

I mentioned in my Afterword to the startle paper that my current research with Robert Levenson is focusing on just these questions. My hunch is that *some* individuals are consistent across emotions in the magnitude of their expressive behavior, and some people are not. I expect also we will find that there are *some* individuals who are consistent in the magnitude of their expressive behavior for just one emotion. We already know from our research in progress that some individuals show coherence in the magnitude of response expressively and subjectively (and I expect physiologically) and others do not show such coherence.

Ruch reported that what he considers a mood—cheerfulness—determines whether one is a hypo- or hyper-expresser in response to humor stimuli. He also found that extroverts showed enjoyment smiling more than introverts, and in his Afterword he reported other work that found three state measures (cheerfulness, low serious, and low bad mood) were related to facial signs of enjoyment. It is not certain from his findings whether these individual differences in enjoyment smiling are due to a trait or state phenomenon. It might be that extroverts more than introverts are in a cheerful, low-serious, low-bad-mood state. It is also possible that even when both introverts and extroverts are cheerful, not serious, and not in a bad mood, the extroverts would show more smiling. Clearly more research is necessary examining subjects with one or another personality configuration when they are in one or another mood state.

I believe that in addition to transient determinants of the magnitude of expressivity such as mood, there may be more enduring individual differences that have a constitutional basis, but that only some individuals will manifest such enduring consistency in the magnitude of emotional expressivity. But we are far from having relevant evidence as yet. (I return to some of these same issues later in the section on applied research, when considering Keltner and Moffit's findings on adolescents who show a high incidence of anger.)

Quite apart from the issue of individual differences, it has been uncertain, from previous literature, whether there ever is, for anyone, a very strong relationship be-

tween what emotion is shown on the face and what emotion is reported. Two chapters make clear that we have underestimated the strength of the relationship between expression and subjective experience. Ruch (chapter 4) showed that a number of artifacts may produce an underestimate of the correlation between expression and self-report. Within subject designs, with aggregated data, yield quite high correlations. Rosenberg and Ekman found the typical procedure of asking people to report on their emotional experience some time after that experience occurred led to an underestimate of the correlation between expression and reported emotion, especially if more than one emotion had been experienced. The procedure they employed, which provided a means to help subjects retrieve memories for their specific emotional experience at specific points in time, yielded strong relationships between expression and subjective report. It remains to be determined whether substantial individual differences in the coherence between expression and report will still be found when the procedures used by either Ruch or Rosenberg and Ekman are adopted.

Keltner's paper strongly suggests that there is a distinctive signal for embarrassment. Although he describes what he terms "smile controls" early in his chapter as including a variety of behaviors (table 6.2), it is clear from that table that most of those facial actions were not unique to embarrassment as compared to amusement. In fact, many of them were more common in amusement than embarrassment. In correspondence, Keltner made clear that he considers only lip pressing (AU 24) as a smile control in embarrassment expressions. Our correspondence also made clear that he agrees with my supposition that observers will call an expression shame as compared to embarrassment if the head and eyes are down, with no visible facial activity (AUs). (In recently reviewing the films I took in 1967 of a preliterate culture in New Guinea, I found that when I asked them to pose shame they all put their heads and eyes down. I did not report this finding before now because it did not involve facial expression per se.) Keltner also agrees with my hunch that if he had allowed his observers in Study 5 the choice of sadness, in addition to shame, they might well have not showed such high agreement about identifying a shame expression.

We have strong evidence that head and gaze downwards is not required for observers to label an expression sadness if the face shows some variation of the facial actions we have proposed for this emotion: inner corners of the brows drawn together and upwards, cheeks raised, slight deepening of the nasolabial fold, and slight depression of the lip corners (AUs 1+4+6+11+15). We don't know what would happen if observers were given a choice among sadness, guilt, and shame when judging such as expression. If the gaze was down and the eyes directed down, but the 1+4+6+11+15 expression was visible, would observers call it sadness, guilt, or shame? It would be best in such a study to allow the subjects to give their own response, rather than to provide them with fixed alternatives. My hunch is that sadness will be the predominant response to such a stimulus, but that needs to be determined.

The embarrassment signal that Keltner has identified differs in kind from all the other facial expressions of emotion. Unlike happiness, anger, fear, sadness, surprise, disgust, and contempt, all of which can be signaled just by facial muscular actions, embarrassment requires changes in gaze direction and head movement and, perhaps also, face touching. Gaze direction and head movement can certainly contribute to the other emotional expressions (gaze and head pointed down contribute to sadness, head turned to the side in disgust, backwards head movement in fear, looking down the nose in

comtempt), but they are not required as they seem to be for the embarrassment signal. Embarrassment also seems to require a sequence of actions, while the other emotional expressions are more like snapshots rather than sequences over time. Another way to consider Keltner's embarrassment findings are to consider this a compound signal, both in the sense of requiring behaviors in additon to facial expression (gaze and head position) and in not having one apex in which all the actions are shown at the same time, but a flow or sequence of actions. It is conceivable that there may be other emotional states that could be constructed in this fashion.

Prkachin also mentioned the importance of examining the sequence of actions in the pain expression, explaining his impression that some actions typically appear early and others late in the expression. It would be important to determine whether these sequential features, in either embarrassment or pain, are in highly overlapping time—for example, with all actions occurring within the same apex period or nonoverlapping time, with some actons beginning after other actions have started their offset or completely offset.

Can one be embarrassed without being either amused or ashamed? If so, it would likely differ from the signal described by Keltner, either not having a smile, or perhaps having one of the nonenjoyment smiles. Craig interprets the smiling activity that occurs in simulated pain displays as embarrassment, but he does not mention any of the other signs that Keltner has observed for embarrassment.

Keltner suggested that mouth opening (either AU 26 or AU 27) is part of the amusement facial expression. I doubt that this is so. Instead, I believe that as positive emotions increase in intensity the mouth opening will increase, just as AU 12 (zygomatic major) increases. (Ruch, in personal communication, 1995, agrees with me.) Laughter can occur in response to a variety of different positive emotions, not just humor, although I suspect the sound of laughter in response to humor is different from the laughter in response to tickling, for example. Laughter, too, can vary in intensity from a chuckle to peals of laughter, and the extent of mouth opening will also vary. Thus, I suggest that one could not tell from the facial expression itself whether the positive response being shown was in response to something humorous, something involving sensory pleasure, strokes to the ego, and so on. The research to determine whether my conjectures are correct has yet to be done.

Craig et al. and Prkachin reported evidence, from their own work and that of others, of a unique facial expression for pain. While no one considers pain to be an emotion, it is important occasion for emotion (in that sense akin to the startle facial expression) and it is noteworthy that it has its own distinctive facial marking. Prkachin accomplished for pain what we have yet to extablish for any of the emotions, but instead presume: the same four actions occur across four very different types of pain stimulation. This is exactly the type of research that is needed to establish firmly the existence of any signal. His findings— generality in the facial pattern across different types of evocations—with Craig et al.'s findings on the ability to simulate and inhibit this expression delineate some of the most important steps that need to be taken in research on emotional expressions. Of course, it is much easier to specify reliable stimuli to elicit either pain or the startle than it is for the emotions, and easier to do so in a laboratory situation. Prkachin speculated that the pain expression may be universal, and emphasizes the need for cross-cultural research. I have never reported my observations of pain in the preliterate culture in New Guinea that I studied in 1967, but I did observe the very configuration that these authors describe when I saw people experience pain.

Craig et al. pointed out in their Afterword the importance of such a signal for people who cannot verbally report on their pain, such as infants, people with intellectual handicaps, brain injuries, and the like. Consistent with this, Prkachin described the primary function of pain behavior in enlisting the aid of others. Paradoxically, such a signal so vital for survival can be, as Craig has shown, very convincingly simulated and inhibited (the pain expression differs from the startle which cannot be either totally supressed or well simulated).

Craig et al. described the pain display in terms I have used to characterize emotional expressions, constituting not one expression, but a theme and variations. Importantly, Craig provided us with some of the determinants of the variations, something which has not been specified for the emotional expressions.

Ruch suggested that exhilaration is the best term to describe the response to humor, tickling, and laughing gas. Would he think amusement is the low end of that state? Exhilaration seems to be on the same intensity level as the emotion terms used to designate other states (anger, fear, sadness, etc.), while amusement appears by definition to be low in intensity, more akin to being annoyed than being angry. While these are semantic matters, they become relevant to research on facial expression in designating what state is being examined, and in how one deals with the self-report of emotion.

Seven of the 11 chapters in this part dealt with the differences between two forms of expression. In five studies the contrast was between expressions that were emotional or not. In the startle and pain studies, it was between the expressions of those states and attempts to inhibit and attempts to simulate those states. Despite this commonality, it is difficult to integrate the findings from all seven studies because each examined facial behavior elicited in a different way. Ekman, Friesen and O'Sullivan dealt only with behavior during an interaction, and that was the situation in Craig et al., and in one of Frank et al.'s conditions, but he also had a condition in which the behavior had been emitted when alone, which was the condition in Ekman, Friesen, and Simons; Gosselin et al.; Hager and Ekman; and Hess and Kleck.

It might seem as if all seven studies examined false or deceptive behavior, but there were differences among them in how this was done. Ekman, Friesen, and O'Sullivan measured masking smiles—smiles in which there is evidence of negative emotions. Frank examined nonmasking, nonenjoymemt smiles. Ekman, Friesen, and Simons and Craig et al. examined a genuine expression (startle or pain), attempts to inhibit that expression, and attempts to simulate it. Hager and Ekman examined simulated startles, posed emotions, and deliberately made single facial actions. Gosselin et al. had actors show emotions without feeling the emotion, and Hess and Kleck had their subjects pose in one study and attempt to deceive in another. Despite these variations everyone found evidence of differences in the facial behavior that occurs when an emotion is presumably experienced and when it is not. There were some consistent differences in the morphology and in the timing or the expressions.

No one dichotomy is sufficient to capture the differences in the two kinds of behavior examined in studies. The distinctions between voluntary-involuntary, deliberate-spontaneous, emotional-simulated emotional, emotional-posed, frank-deceptive, and alone-social are each useful but more than one of them must be considered in specifying just what is being studied.

The condition that I find of least interest is posing an emotional expression on request. It is uninteresting both because this never happens in ordinary life, only in the

psychological laboratory, and because we don't know how people perform this task. It certainly can be done by various means, including, for example, remembering or imagining an emotional situation or remembering an emotional face and imitating it. Posing may sometimes actually generate the emotional experience, and some investigators have used self report as a way to identify whether that did or did not happen.[1] Acting, of course, does occur in real life, but we do not know the extent to which actors do the same things that nonactors do when they are trying to convince others that they feel an emotion they do not feel. When nonactors try to convincingly simulate an emotion, I call this lying, not acting. The crucial difference is that the actor notifies his target that it is performance, while the liar does not (Ekman, 1985).

Looking across these seven studies, and adding some speculations to fill in a few gaps, we find that there are some promising findings on how to distinguish between two types of facial movement, most of which fit Ekman and Friesen's (1982) and Ekman's (1985) forecast. I will compare what occurs on the face when a person (1) spontaneously shows an emotional expression or a pain or startle expression, when there is independent evidence (subjective report or physiology) that either an emotion(s) or startle or pain is being experienced; and (2) deliberately makes a facial movement for the purpose of intentionally misleading another person into believing an emotion or pain or a startle is being experienced when it is not, without giving notification of that intention. I don't think it should matter whether the person is alone or with another person in either of these conditions. Also, I believe that the characteristics I am about to describe for category 1 would also be manifest in other (beyond the startle) reflexive facial actions, or for nonemotional facial movements that are highly learned habits, occurring involuntarily (e.g., agreement signs and other conversational signals).

Category 1 can be distinguished from category 2 in terms of the following characteristics.

1. *Morphology.* This is, so far, best documented for enjoyment, but I have described (Ekman, 1985/1992, 1993) how the absence of muscle movements that are difficult to make voluntarily could differentiate category 1 from category 2 facial movements. This does not appear to be the case for the simulated pain face that does not differ in morphology from the actual pain face.

2. *Symmetry.* While difficult to measure, this is potentially a very valuable differentiator, with symmetry being more common in category 1 than in category 2.

3. *Total duration.* The very brief (under 2 seconds) and very long (more than 5 seconds) should occur more with category 2 than category 1. Micro expressions (those that last 1/3 sec. or less) are excluded, since they are spontaneous, suppressed, or repressed emotions (Ekman, 1985). There also will be some long-duration expressions that fit category 1, but these will be exceptions.

4. *Speed of onset.* Although this will vary with social circumstance, in comparable circumstances, onsets will be shorter in category 1 than in category 2.

5. *Coordination in apexes.* The apexes of each of the actions involved in the expression will be more likely to overlap in category 1 than in category 2.

6. *Ballistic trajectory.* The expression will appear smooth over its course in category 1, rather than jagged or stepped in category 2, as was found by Hess an Kleck.

We do not know how accurately expressions can be classified as category 1 versus category 2 when all six of these indices are used, but I suspect it will be quite high.

Nor do we know whether some of these indices are more powerful discriminators than others, for all people or for some people. (I suspect that no. 2 and no. 3 will prove less powerful than the others in distinguishing category 1 and category 2.) Again, a profitable area for further research.

The research findings reported in this section and in other sections of this book (see especially chapters 15, 16, 17, and 21) strongly suggest that those studying facial expression need to measure at least some of the six parameters listed above if they want to ensure that they are studying emotional rather than nonemotional expressions. The simple fact that people will say that they felt an emotion does not mean that every expression that preceded it was emotional. And because of demand characteristics, sometimes people will report feeling an emotion when in fact they did not. While I am not suggesting that the face is the "gold standard" for knowing when an emotion occurs, one can increase the likelihood that an emotion is being examined if one utilizes at least some of the six parameters of facial movement described above to distinguish the emotional from the nonemotional facial activity.

The evidence reported in this book and other studies summarized by Frank et al. on smiling make overwhelming the need to apply at least some of these six parameters in the study of smiling. Regrettably some investigators ignore this work and continue to simply measure zygomatic major activity without regard for any of the other six parameters listed here (e.g., Fridlund, 1991).

A number of the authors interpreted their findings—emotional expressions occurring when their subjects thought they were alone—as contradicting Fridlund's (1992) claim that emotional expressions do not occur except in the presence of others. Of course, Fridlund would likely reply that people who show facial expressions when alone are imagining that other people are present. That line of reasoning makes Fridlund's position not subject to disconfirmation and therefore of little use. Fridlund (1994) more recently has argued that facial expressions are unrelated to emotion, that they tell us about motivations and intentions, not internal states. I (Ekman, 1989; in press) have argued that this if a false and misleading dichotomy. An evolutionary view of emotion regards the internal changes during emotion as preparation for responses, most often to others, that have been adaptive and messages about motivations and action intentions to be among the kinds of information that can be conveyed to others by an emotional expression.

Elsewhere (Ekman, 1972, 1989, 1992, in press) I have suggested that the presence of others can act to either amplify or deamplify facial expressions of emotion, depending on the role relationship between the people, the emotions being felt, and other aspects of the social context. It is far too simple to maintain that the presence of others will have but one uniform effect on expressiveness. The fact that people experience emotions when alone (e.g., in response to a sunset, lightning, etc.) does not diminish my claim, and that of other theorists who operate from an evolutionary perspective, that emotions evolved to deal with fundamental life tasks that involve other people, such as mating, fighting, or caring for children.

The two developmental studies in this section contribute important information on fundamental issues about emotion and expression. Both studies show that there is differentiation in facial response, in Rosenstein and Oster, in the first two hours of life, and in Carmas et al., in both 5- and 12-month-old infants. To show that expressions are differentiated does not mean that the specific facial expressions that signal emotion in

adults are evident so early in life. Both papers report evidence that strongly suggests that is not so, that there may well be infant facial expressions of emotion that are different from adult facial expressions of emotion. Studying their results, it does not appear that these infant facial expressions of emotion are unique, bearing no relationship to their adult counterparts, but they are far from identical.

Camras et al. provided evidence on another fundamental issue—the universality of facial expression. Their comparison of Japanese and American infants provides uniquely important evidence for universality. It is important because it is the first evidence on infants and also because it adds to the small amount of evidence on universality for the display of emotions, not just their recognition. We now know that Japanese and American *infants* do display the same emotional expressions to the same situation. The first evidence of comparable universal facial expressions in Japanese and American *adults* was obtained 25 years ago in our (Friesen and my study, reported in Ekman, 1972) study of college students. That work used the predecessor of FACS, the Facial Affect Scoring Technique (FAST), and found that when Japanese and Americans were each alone watching a slightly pleasant and a very unpleasant film, very similar facial activity occurred.

Camras et al. did find a cultural difference in the latency of negative emotional expressions, with Americans responding more quickly than Japanese to the arm-restraint procedure. They pointed out that, contrary to most theories that cultural differences would increase over time, it was just the opposite. This difference was apparent only in the 5-month-old infants, not the 1-year-old infants. Their explanation seems sensible, and presumably they will be able to replicate this finding in the study described in their Afterword.

Rosenstein and Oster raised the question of why the infant shows a particular expression in response to particular stimuli—what they call the mouth gaping in response to a bitter taste and lip pursing in response to a sour taste. Few investigators since Darwin have raised such questions, and their explanation is quite consistent with how Darwin reasoned, although Darwin did not know of this expression. They show that detailed measurement with Baby FACS, using an empirical approach (as opposed to an approach based on a priori emotion classifications, found in Izard's work), reveals information about the sensory capacities and hedonic responses of newborns that other response measures failed to find (and other facial coding systems would have misinterpreted or missed).

Oster, in her Afterword, raises a very important point about the hazards of using a priori formulas, particularly in infants, to interpret what appear to be "blends," in which one might want to presume that there are two different emotions occurring at once. Oster raises the question of whether elements of AU 9 and AU 12 should, in the infant, be interpreted as disgust and enjoyment. She has instances that suggest that this is not the case, and I have reason to believe she is right not only about infants but about adults as well. Part of the problem is that certain AUs may have different meaning in different contexts, or when shown by infants as compared to adults. Part of the problem is being sufficiently careful to ensure that an AU is actually present, when only one of the appearance changes associated with that AU occurs, rather than all of those changes.[2]

Rosenstein and Oster also note that AU 18 occurs with interest or puzzlement. I have observed this same action in some adults as a mark of these same states.

Applied Research Issues

Six of the eight chapters in this part dealt with depression and/or schizophrenia. The results are not entirely consistent, but that would not be expected given the disagreements about diagnosis that are likely when research is done in three different countries and over a 20-year period. To further complicate comparisons across the studies, some studies made subdistinctions (blunted and nonblunted schizophrenics; in or outpatients; suicidal or nonsuicidal), which other studies did not.

There were differences in the occurrence of enjoyment smiles and nonenjoyment smiles for depressed and schizophrenic patients as compared to controls (Berenbaum and Oltmanns; chapter 16), for schizophrenics as compared to controls (Steimer-Krause et al., chapter 17), and in relationship to clinical improvement for depressed patients (Ekman et al.; chapter 15). Contempt expressions were more evident in schizophrenics than controls (Steimer-Krause et al.; chapter 17), in depressed patients who do not show clinical improvement (Ekman et al.; chapter 15), and in suicidal as compared to nonsuicidal patients (Heller and Haynal; chapter 19). Two studies suggest that patients show less use of the face to illustrate speech—what I termed (Ekman, 1979) *conversational signals,* than controls (Steimer-Krause et al., for schizophrenics as compared to controls; Ellgring for depressed as compared to controls). Heller and Haynal's finding of a diminution of upperface activity for suicidal as compared to nonsuicidal patients may also reflect a decrease in conversational signals. There was also evidence that measures of facial expression can predict subsequent behavior, (ratings of clinical improvement, Ekman et al.) and outcome of psychotherapy.

It is difficult to integrate Ellgring's report with the findings of the others who have studied psychopathology, because Ellgring's use of FACS is so different. His was a very early study, done before Friesen and myself had told those using FACS how to integrate separate facial actions into combinations or "events" when the different facial actions occur in overlapping time. Facial expressions often involve the actions of two, three, four, or five actions, which may start at slightly different moments and time, but merge together into a configuration that is held for a few moments before fading off the face. FACS provides the investigator with a means to identify each of those separate actions (AUs) that produce any configuration, but proper analysis requires that the investigator reassemble them in the original configurations. Ellgring did not do so. He treated each AU as a separate event, except for combining the two actions that combine to raise the entire eyebrow (1+2). The failure to reassemble the separate AUs into configurations, and then analyze those configurations, is responsible, I believe, for why Ellgring found primarily individual differences rather than many common features in what expressions changed over the course of depressive illness. Another problem is that Ellgring did not have available the "emotion dictionary" that Friesen and I had developed to interpret those configurations, translating them into emotion scores. There are many different configurations for any single emotion. If Ellgring had reassembled the separate AUs and then used the interpretative dictionary to convert those AU scores into emotions, I expect he would have found much more commonality across patients, as did the other investigators who studied psychopathology and who took those steps. Ellgring's work also was done before the distinction between enjoyment and nonenjoyment smiles was published. ans so we cannot know if the increase he noted was in one or both forms of smiling.

Berenbaum and Oltmanns demonstrated that facial expression measurement provided information not available from other sources. Blunted-affect schizophrenic patients were less facially responsive that nonblunted schizophrenics, but there was no difference in reported emotional expreience. They interpreted this as showing that blunted affect is a disturbance of expression but not of feeling. Of course, it is possible that the self-reports were not valid, but reflected demand characteristics. While it is difficult to know whether or not that is so, their study does tell us that facial measurement provides information not available in self-report: facial behavior differentiated blunted affect from nonblunted-affect patients, self-report did not. Their study also showed that measures of emotional reactivity based on standard clinical interviews were not sufficient to discover the ways in which psychiatric groups differ from each other and from normals, while measures of facial behavior made those distinctions. This is consistent with Ekman et al.'s finding that clinical ratings on the Brief Psychiatric Rating Scale (BPRS) did not improve predictions of inprovement made on the basis of facial behavior. (Later, in discussing methodological issues, I describe other studies that found that facial measurement provided information not available from observers judgments; see chapters 10, 14, and 21.)

Steimer-Krause et al. found more facial activity during talking, except for the psychosomatic patients. Ellgring reported more facial activity during speaking for normal subjects. Presumably these findings are due to the fact that most of the conversational signals occur when speaking to emphasize speech as it is spoken. Although there are also listener facial responses, these typically do not occur as often as the conversational signals used by the speaker. In addition to conversational signals, I expect that in general there are more facial expressions of emotion when speaking than when listening, although there will be exceptions—for example, when someone is telling you terrible or wonderful news.

Steimer-Krause et al. also make the very important point what when we see an emotion on the face we do not know from the expression itself who is the target of the emotion. For example, the contempt they found that differentiated schizophrenics from other patients or controls may have been directed at themselves, at others not present, or at the person with whom they are interacting. To discover the target of the emotion shown on the face one must examine what is being said and gaze direction. Krause, in his Afterword, cited another study by their group that found that schizophrenics misinterpret who is the target of the emotions shown by others. The schizophrenics interpreted the enjoyment smiles shown by others as a sign that the other person is enjoying the schizophrenic's misfortune. Baenninger-Huber (who has worked with the Krause group) made a similar point, adding that not only do we need to consider what is said or not said when a facial expresion of emotion occurs, but we also need to attend to the sound quality, in terms of what emotion is being signaled vocally.

Baenninger-Huber described a number of very intriguing concepts illustrated by her study of a single patient during treatment. Individuals regulate their affect, she said, by evoking through their own behavior specific reactions, emotions, and fantasies in the persons with whom they are interacting. In correspondence she told me this idea was first forwarded in another paper (Baenninger-Huber, Moser, & Steiner, 1990), and is similar to ideas forwarded by Krause. In German, she wrote, the word for this is *Ubertragungsangebot*, which comes from psychoanalysis and can be translated as a transference offer.

This formulation provides an interpretation of *why* individuals display a particular array of emotional behavior—it is to reenact with another an particular interaction pattern, presumable one that repeats elements of their own interaction pattern with their parents. Although her data cannot prove this interpretation, her work and other papers by the Krause group importantly attempt to explain individual differences in the pattern of emotional behavior.

Baenninger-Huber's formulation is consistent with how Keltner and Moffit interpret their finding that externalizing adolescents show more anger. They point out that the high display of anger expressions is likely to evoke negative emotions in others, which in turn contributes to the social rejection and hostility that such adolescents experience from others. They do not, however, go as far as Baenninger-Huber. They do not say that these adolescents are motivated to act this way because they need to cause others to reject them, repeating presumable traumatic or unresolved relationships with their parents.

Clearly an explanation is needed for why people would act in a way that causes others to act unpleasantly toward them, whether adolescents, as in Keltner and Moffit's study or adults who glare, as in the study by Chesney et al. Other alternatives are that people who show a great deal of anger do so because earlier (traumatic?) experience has affected them in a way that: (1) has caused anger to dominate their affective system in a monopolistic fashion, generally being the first emotion to occur in response to a variety of appraisals; (2) has lowered the threshold for anger being expressed; (3) has damaged their capability to inhibit anger expressions; or (4) to use Lazarus's (1991) terminology, has caused the knowledge they draw upon in their primary appraisals to lead them to repeatedly interpret a variety of events as demeaning or personally assaultive.

These are not mutually exclusive explanations. Some may apply to one person, other explanations to other people. Or it may be that all of them apply to one person, but some of these explanations may be more relevant at one point in time than another. The important matter is that we are beginning now to establish marked individual differences in facial expressions of emotion, differences in the predominant emotion or emotions that are shown in certain circumstances, differences in the magnitude of the facial expressions, and perhaps also differences in the latency of the response. Much more work is needed to establish the stability of those individual differences, their developmental course, and how these differences in facial expression relate to differences in vocal expression. speech, and coincident physiological activity. Emotions are not just transient phenomena. There are stable individual differences in characteristic facial expressions of emotion. Now we can begin to design research not only to better document the nature of these differences but to begin to allow us to choose among various explanations of those differences.

Chesney et al. raise the interesting possibility that there may be signs of personality in facial activity, not just signs of emotion. The "glare" configuration, which they found more often in people presumed to be at risk for coronary artery disease than those presumed not to be at risk, *might* be such a sign. The actions that compose the "glare," they explain, could be controlled anger, low-intensity anger, or perhaps a marker of the personality trait of hostility, not the emotion anger. They cite indirect evidence that suggests it is most likely that glare is not an emotion sign but a sign of hostility, but the matter is far from settled. Ekman (1985) earlier suggested that accurate

inferences about personality can be drawn by observing when specific emotions are shown, how often they are shown, and how strongly they are shown. So, for example, we think a person is hostile if that individual shows more anger than others, in circumstances others do not show anger, and when the person shows anger it is more intense and longer lasting than it is for others. But the Chesney et al. study raises an entirely different possibility—that there are direct signs of personality, perhaps related to emotion signs, but unique to personality.

Methodological Matters

The chapters in this book also illustrate a number of important methodological issues. As mentioned earlier, both the chapter by Rosenberg and Ekman and that of Ruch show that most prior research has probably underestimated the extent of the correspondence between facial expressions of emotion and subjective experience. Each of these chapters describes a very different methodological solution to the problems that have led to an underestimation of that correspondence. Although their solutions are not applicable in every research design, they should be considered by anyone intending to measure both expression and self-report of emotion.

In the study of smiles when lying, Ekman, Friesen, and O'Sullivan used EMFACS rather than FACS. By doing so, they reduced the scoring time, but because they used EMFACS they could not discover any other facial signs of lying or truthfulness that are not, strictly speaking, emotional behaviors. EMFACS focuses just on those facial actions for which there already is good evidence of their relevance to emotion. Other facial actions that might signal emotion, or other facial actions which are not emotional but might relate to the phenomena under study, are ignored by EMFACS. For example, the glare facial expression identified by Chesney et al. would not have been discovered if they had used EMFACS instead of FACS. While we have subsequently modified EMFACS to include the AUs relevant to glare as a result of this study, there may be other such expressions that will not be discovered if EMFACS is used rather than FACS. The use of EMFACS also makes it impossible to evaluate all of the other signs of emotional behavior apart from morphology listed above. Rosenstein and Oster's study is another case in which if an abbreviated a priori set of emotion categories had been scored, the differential responses to negative tastes they uncovered with Baby FACS would not have been found.

Ekman et al. showed that EMFACS misses some of the facial expressions that FACS is able to identify. This should not be a surprise, since to economize on time EMFACS does not allow slowed motion replaying. Again to economize on time, EMFACS does not score brow raises (AUs 1+2) and brow lowers (AU 4), when these occur alone. Often these actions are conversational signals, indicating emphasis, determination, or perplexity in the speaker, or an exclamation or perplexity in the listener. Two of the chapters reported differences between psychopathological groups and normals in these actions, and those findings would not have been obtained if EMFACS had been used.

To argue the opposite viewpoint, there is no reason to do a full FACS scoring unless there is a theoretical or methodological need to do so. In particular, the measurement of duration rather than frequency, which is very labor intensive, should be re-

served for the following situations: when it is necessary to identify exactly places where emotion occurred so that other phenomena, such as physiological activity, can be examined at those moments (e.g., Davidson, Ekman, Saron, Senulis, & Friesen, 1990); when there is a hypothesis about duration itself, such as Frank et al. examined regarding the difference between enjoyment and nonenjoyment smiles; and when there is reason to believe that at least some subjects will show infrequent but long-duration facial expressions. This last reason for measuring duration refers to unusual occasions, for typically there is a very high correlation between frequency and duration.

To further measure the duration of onset to apex, the apex duration. and end of apex to offset takes additional time beyond simply getting the duration measure from onset to offset. Some of the chapters in this part, however, illustrate when precisely this type of information is needed (see Ekman, Friesens, and Simon; Frank et al.; Hess and Kleck; and Ruch's article on alcohol). Prkachin's study shows that duration and intensity measures were highly correlated; however, some of his findings were better revealed with one measure as compared to another. In the initial study of any phenomenon, it would appear wise to measure intensity, frequency, and duration, and perhaps some of the onset, apex, and offset measures. As more is known about a type of expression, it may become apparent, as it is now with the pain and startle expression, which of these indices can be eliminated.

Berenbaum and Oltmanns appear to have taken an opposite position, recommending the use of FACS only when the investigator has a very specific hypothesis to test. While I still maintain FACS is the technique to use rather than EMFACS when discovery is the object, and in the initial stage of research, perhaps Berenbaum and Oltmanns meant that FACS can be used selectively to test a hypothesis about a specific emotion. If, for example, one is only interested in assessing changes in fear, then FACS could be used to score just fear-relevant AUs, obtaining all of the timing measures that FACS entails but ignoring all AUs not relevant to fear.

If an investigator expects to collapse all of the FACS scoring into positive versus negative emotional expressions, as did Berenbaum and Oltmanns, then they might as well use EMFACS rather than FACS. Similarly, if all an investigator intends to do is to use a single summary score of the occurrence of each emotion in each subject or patient, then EMFACS may be preferable. If the signs of emotion are likely to be subtle, slight, fleeting, or masked, as commonly occurs in deception research, then again FACS is the method of choice.

Gosselin et al. utilized the tables published in FACS (Ekman & Friesen, 1978) to interpret which facial actions signify which emotions. Those tables were meant only as a beginning point to describe what we then thought were the prototypic combination of Action Units (AUs) for each emotion. They are by no means comprehensive and leave out many, perhaps the majority, of the combinations of AUs that, thought not prototypic, have been found to be signs of each emotion. Currently there are two methods for interpreting the emotional significance of AUs. Friesen and Ekman's (1989) "emotion dictionary" makes interpretations of AUs based on theory, findings, and intuitions. It was developed by specifying which AUs are required to signal a particular emotion, which other AUs are irrelevant, and which change the emotion interpretation. The dictionary then extrapolates logically to many other combinations of AUs. Ekman, Irwin, Rosenberg, and Hager's more recent FACS Affect Interpretation Data Base (FACSAID) accomplishes the same task in a different fashion. It was derived from the

authors' examination of each of more than 5000 AU combinations that occurred in a large data set. Interpretations were based, again, on theory, findings, and intuitions. The database has grown as new AU combinations have been encountered in other data sets. It is likely that the interpretations made on the basis of the "emotion dictionary" and FACSAID would be the same, but there has as yet been no empirical test of that.

A number of chapters in this book illustrate the difference in the information obtained when results are analyzed on the purely descriptive AU level and when AUs are grouped on the basis of what emotion they signify. Heller and Haynal found nothing happened more than once when they performed their analysis just on the basis of AUs, but clear differences emerged when they analyzed their results in terms of the emotions signified by those various AUs. Steimer-Krause et al. provided an especially good illustration of analyses conducted at different levels, from the purely descriptive to the interpretative.

The studies by Frank et al., by Keltner, by Rosenstein and Oster, and by Prkachin illustrate the difference in the type of information that can be obtained by measuring the face as compared to having observers make judgments when viewing faces. They show also the value of using both approaches. The only way to test some of Ekman and Friesen's ideas about the characteristics that distinguish an emotional from a nonemotional smile was for Frank et al. to actually measure facial behavior. It is quite a separate question as to whether or not observers can accurately differentiate one smile from another. If they had just done a judgment study, they would have found that accuracy is possible, though meager, but they would not have been able to establish what are the actual signs in the face that distinguish one kind of smile from another, nor could they have discovered whether the differences between smiles are robust or modest.

Keltner's work to establish a signal for embarrassment meant that he had to do both a measurement study to discover what are the signs of embarrassment and a judgment study to find out whether people do recognize embarrassment and distinguish it from related states. It is conceivable that there could be a reliable sign of a state (such as the enjoyment smile versus nonenjoyment smile) that is not a very robust signal. Keltner's initial judgment study also obtained quite modest accuracy, but when he used his results from the measurement study to select expressions that had the various elements he had uncovered in the measurement study, he obtained higher agreement among observers in distinguishing embarrassment from other emotions.

Prkachin reports in his Afterword that facial measurement has outperformed naive observers' ratings of pain. He notes that the information is there in the face, but observers do not know what to attend to. He reports taking the obvious next step to train observers to be more accurate, but did not tell us the findings.

Rosenstein and Oster show that measurements could discriminate which stimuli the infants were responding to, while observers could not make some of these same discriminations. If an investigator were to only use observers' ratings as the tool for measuring the face, the investigator might well think that differences that Rosenstein and Oster found do not exist. Whenever observers fail to make a discrimination, it does not mean that the difference is not there, only that the observers don't recognize it. Similarly, as described earlier, both Berenbaum and Oltmanns and Ekman, Matsumoto, and Friesen found differences in studies of psychiatric groups that ratings made by observers did not distinguish. To show that a difference is or is not present requires the use of FACS (or if it is infants, Baby FACS).

This is an appropriate place to applaud Oster for her work in developing and promulgating this adaptation of FACS. It is true to the spirit of FACS in providing a means of closely describing facial activity, taking account of anatomical differences between the infant and adult behavior. It is consistent with FACS, and unlike MAX and Affex, in separating description from inference about emotion and in being comprehensive and not selective. Oster also provides descriptive codes for a variety of facial actions encountered in the infants. And she has separated brow lowering from bringing the brows together, utilizing AU 3 as well as AU 4. We originally had provided AU 3 for just that purpose, but dropped the distinction worried that it would not be reliable. I am convinced now, from Oster's work and our own observations, that the distinction can be made reliably.

In her Afterword Hess described a subsequent study in which she and Kleck had observers judge the different kinds of facial expressions she had measured in her chapter of this book. Noting the modest level of accuracy achieved by judges in Frank et al. and in her follow-up study, she raised the question as to why observers are not more accurate than they are. She suggested that observers follow a decoding display rule, trying to interpret the situation in the way it is intended and, presumably because of this, disattending to the clues that would allow them to better identify enjoyment from nonenjoyment smiles. We (Ekman, 1985/1992; Ekman, Friesen, & O'Sullivan, chapter 9) suggested that this is characteristic of deceptive encounters. Lies often succeed because the target of the lie is collusively involved, supporting the lie and ignoring discrepant information.

Ekman, Friesen and O'Sullivan's chapter raises the question of whether highly motivated people might be more accurate than others in distinguishing emotional from masking smiles. Ekman and O'Sullivan's (1991) subsequent study of such motivated people (law enforcement and national security personnel) examined accuracy in detecting deceit, not just accuracy in differentiating two types of smiles. They found that even such highly motivated people are not very accurate in judging who is lying and who is truthful. It remains to be known whether it is possible to train people to make these judgments more accurately.

Prkachin's article is notable for the excellent use of graphic displays to help in understanding the occurrence of particular AUs in the different types of pain stimulation. Baenninger-Huber also used graphic displays in an interesting way. Prkachin's work is also notable in the use of factor analysis to aid in identifying which AUs are consistently displayed with pain. The use of a drawing to illustrate the pain face is quite successful also, as were the photographic illustrations Keltner used for embarrassment.

Ellgring pointed out the need to develop better training materials for FACS. We wrote the training manual and materials nearly 20 years ago. We have changed how we use FACS, but those changes are provided to those who have learned FACS in a separate document, only after they have learned FACS. If we (Friesen and myself) do not rewrite the manual to incorporate those changes, the next generaton of facial researchers certainly should do so. They should consider making use of the new developments in interactive video, which could well speed up how long it takes to learn FACS.

Another important matter is to provide the person who has learned FACS additional training in its use. Rosenberg is working, with my advice, on this. And, on the horizon, perhaps less than five years away, are automated techniques for measuring visible facial movement, which may not totally eliminate the human who scores the

face but will significantly reduce the time it takes (for a recent review of such work, see Ekman, Huang, Sejnowski, & Hager, 1994).

Conclusion

There can be no single conclusion from this set of very interesting studies beyond the very simple and dramatically illustrated fact that measurement of facial behavior reveals unique and important information about a wide variety of psychological phenomena. The richness and complexity of facial behavior cannot be captured by a still photograph, nor can the information that facial activity reveals be assayed by asking an observer to chose one from a short list of emotions or to use a rating scale. The tools for mining the flow of facial activity are available; the work required is arduous but not overwhelming; and as these chapters show, the rewards can be very high.

Acknowledgment

Notes

1. I might note that some authors (e.g., Hess and Kleck; chapter 12) use the term "posing" to describe trying to mislead another person, but I think we should reserve that term for when people respond to an investigator's request to produce on command a facial expression. For this reason I do not call the technique that I have used in my joint research with Levenson (Ekman, Levenson, & Friesen, 19834) posing, since in our directed facial action task we gave the subjects muscle-by-muscle instructions to construct an expression without ever mentioning the emotion itself. This is certainly not real life, any more than posing is, but it is a different instructional task than posing.

2. Many adults, especially women, show some signs of wrinkling along the sides of the nose when there is a moderate to strong AU 12, especially if the muscle orbiting the eyes (AU 6) is involved. I do not believe that this is actually due to the action of levator labii superioris, alaquae nasi, which is what AU 9 is supposed to measure. In my observation when this nose wrinkling occurs with a 6+12, it is not accompanied by changes in the upper lip, nostrils, or tip of the nose, which are marks of the action of this muscle. Often this wrinkling occurs as a momentary action, emphasizing the enjoyment signaled by AU 12. Sometimes it appears throughout the AU 12. My observations suggest that it is a consistent feature in some people; they will either never or always show this nose wrinkling with AU 6+12. I recommend, therefore, that we no longer score AU 9 if there is only wrinkling on the sides of the nose without the accompanying changes in the upper lip, tip of the nose, and nostrils. This is a change from the FACS manual, which allows the scoring a AU 9 if the only change observed is more than surface wrinkling on the sides of the nose. I believe this is consistent with how Oster scores 9 in the presence of 6+12 in infants. One more matter: when AU 9 does actually occur with AUs 6+12, should we interpret that as a blend of disgust and enjoyment, as we have been doing in the past? I have my doubts about that; certainly I think one can enjoy the feeling of disgust, but the occasions in which I saw a strong 9 with a strong 6+12 I do not think disgust was being experienced. I have nothing to recommend other than caution in interpreting this combination of AUs.

References

Baenninger-Huber, E., Moser, U., & Steiner, F. (1990). Mikroanalytische Untersuchung affectiver Regulierungsprozesse in Paar-Interaktionen. *Zietschrift fuer Klinische Psychologie, 19*, 123–143.

Davidson, R. J., Ekman, P., Saron, C., Senulis, J., & Friesen, W. V. (1990). Emotional expression and brain physiology I: Approach/withdrawal and cerebral asymmetry. *Journal of Personality Social Psychology, 58*, 330–341.

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

Ekman. P. (1979). About brows: Emotional and conversational signals. In *Human Ethology*, M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog, (Eds.) Cambridge: Cambridge University Press, pp. 169–248.

Ekman, P. (1985/1992). *Telling lies: Clues to Deceit in the Marketplace, Marriage, and Politics, Second edition.* New York, W. W. Norton. (2nd Edition 1992)

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

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*, 169–200.

Ekman, P. (1993). Facial expression of emotion. *American Psychologist, 48*, 384–392.

Ekman, P. (in press). Expression or communication about emotion. In N. Segal, G. E. Weisfeld, C. C. Weisfeld (Eds) *Genetic, ethological, and evolutionary perspectives on human development.* Washington DC: APA Books.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiles. *Journal of Nonverbal Behavior, 6* (4), 238–252.

Ekman, P., & Friesen, W. V. (1989). *Emotion dictionary.* Unpublished manuscript, University of California, San Francisco.

Ekman, P., Huang, T. S. , Sejnowski, T., and Hager, J. C. (1993). Final report to NSF of the planning workshop on facial expression understanding, July 30 to August 1.

Ekman, P., Irwin, W., Rosenberg, E. R. , & Hager, J. C. (1995). *FACS Affect Interpretation Data Base.* Computer data base, University of California, San Francisco.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes between emotions. *Science, 221*, 1208–1210.

Ekman. P., & O'Sullivan, M. (1991). Who can catch a liar. *American Psychologist, 46*, 913–920.

Fridlund, A. J. (1991). Sociality of solitary smiling: potentiation by an implicit audience. *Journal of Personality and Social Psychology, 60*, 229–240.

Fridlund, A. J. (1992). The behavioral ecology and sociality of human faces. In M. S. Clark (Ed.), *Review of personality and social psychology* (Vol 13; pp. 90–121). Newbury Park, CA: Sage.

Fridlund, A. J. (1994). *Human facial expression.* San Diego, CA: Academic Press.

Izard, C. E. (1979). *The maximally discriminative facial movement coding system (MAX).* Unpublished manuscript. Available from Instructional Resource Center, University of Delaware, Newark, DE.

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