Journal of Personality and Social Psychology
1974, Vol. 29, No. 3, 288-298

DETECTING DECEPTION FROM THE BODY OR FACE [1]

PAUL EKMAN [2]
University of California, San Francisco

WALLACE V. FRIESEN
Langley Porter Neuropsychiatric Institute, San Francisco

The experiment was designed to test two hypotheses concerning differences between the face and body when a person is engaged in deception. Subjects were required to be honest in one interview, frankly describing their feelings about a pleasant film, and to be deceptive in another interview, concealing negative affect aroused by an unpleasant film and simulating pleasant feelings. As predicted by the first hypothesis, the face was mentioned more often than the body when the subjects were asked afterward what behavior should be censored or controlled in perpetrating deception. Videotapes of the facial and body behavior during the honest and deceptive interviews were shown to separate groups of observers. The second hypothesis—that when deceptive behavior was judged, more accurate judgments would be made from the body than from the face, but that when honest behavior was judged, there would be little difference in the accuracy achieved from the face or body—was partially supported.

Recent interest in body movement and facial expression among behavioral scientists (cf. reviews by Duncan, 1969; Ekman, Friesen, & Ellsworth, 1972; Harrison, 1973; Knapp, 1972) and in the popular press (e.g., Fast, 1970) stems in part from the belief that nonverbal behavior reveals how people feel, even when they wish to conceal their feelings. The idea that nonverbal expressions of emotion are not as easily censored or disguised as the content of speech is an old one (cf. Darwin, 1955; Freud, 1959) and a recurrent one (cf. Feldman, 1959; Goffman, 1959; Ruesch & Kees, 1956). Despite this long history, there have been no data to support such claims, nor has there been thorough description or theoretical explanation of how nonverbal behavior might escape efforts to censor interpersonal communication.

Ekman and Friesen (1969) recently proposed that certain aspects of nonverbal behavior might escape censorship or serve to maintain deception more than others. They proposed that when an individual is engaged in deception, his body, more than his face, is a source of *leakage* (i.e., the nonverbal act reveals a message otherwise being concealed) and *deception clues* (i.e., the nonverbal act suggests that deception is occurring but does not reveal the concealed message). This article reports a partial test of that proposition.

Ekman and Friesen reasoned that for both neuroanatomical and sociocultural reasons most individuals in Western cultures grow up subject to more commentary, instruction, and reinforcement on their facial activity than on their body movement during conversation. This history of greater social reinforcement for facial behavior than for body behavior, the greater social accountability for what is shown on the face, results in greater awareness of ongoing facial activity and better retrieval for the purpose of simulation. Although people could lie with the body as with the face, Ekman and Friesen claimed they do so less frequently.

Hypothesis 1. When asked what behavior they thought of censoring or employing in simulation, subjects who have just engaged in a deceptive interaction will mention the face more often than the body.

Ekman and Friesen reasoned that as a result of the greater focus on the face than the body, when people attempt to conceal actual feelings and to simulate emotions not felt, they disguise the face more than the

[1] The research was supported by Research Scientist's Development Award 5-K02-MH06092 and by Research Grant MH 11976-06 from the National Institute of Mental Health. We are indebted to Patricia Garlan for her editorial help.

[2] Requests for reprints should be sent to Paul Ekman, University of California, 1405 Fourth Avenue, San Francisco, California 94143.

body. In such circumstances, facial behavior contains lies of omission and of commission. Through careful monitoring, attempts are made to inhibit, interrupt, or mask facial expressions of actual feelings. With more or less skill, attempts are made to convincingly simulate facial expressions and feelings not experienced. Although the body could be similarly disguised, people usually fail to do so. If any attempts are made to control the body, they are typically lies of omission, not lies of commission. Ekman and Friesen noted the complication that within the face, certain expressions (namely, microexpressions, time-reduced remnants of interrupted or inhibited facial muscular movements), may, like the body, provide leakage; nevertheless, they expected that the usual observer misses or disregards these microexpressions and instead is misled by the face. The body, however, usually more truthfully reveals to the observer either how the person actually feels (leakage) or the fact that something is amiss (deception clues). This difference between face and body would be limited, however, to situations in which the person is engaged in deception. When there is no deception, when communication is frank, then there should be little difference in the information provided from the face and body.

Hypothesis 2. When observers are required to judge whether a person is honest or deceptive, more accurate judgments will be made from the body than from the face when deceptive behavior is considered; when honest behavior is considered, there will be little difference in accuracy between judgments of the face and body.

Pilot studies suggested that there are marked idiosyncracies in a person's repertoire of both facial and body behavior when no attempt has been made experimentally to arouse intense affect. Without knowledge of these idiosyncracies, observers might misconstrue the behavior when judging whether a person is deceptive. This consideration led to the decision to test Hypothesis 2 with groups of observers who were given some acquaintance with each person's nonverbal behavior, as well as with groups who had no such additional information.

METHOD

Deception Situation

An interaction in which deception was central was devised, based on Ekman and Friesen's (1969) description of the four dimensions which distinguish deceptive interactions: the saliency of the deception, the stakes for success, the balance of roles, and the extent of antagonism between deceiver and deceived about the maintenance of deception. Saliency was achieved by instructing both interviewer and subject that in any one of their four interviews, the subject might be required to deceive the interviewer. Antagonism was created by informing the subject and the interviewer that the subject was to attempt to mislead the interviewer, if so required, while the interviewer's aim was to uncover any such action. High stakes for success in deception were attempted by (*a*) telling the subjects that success in deception was relevant to their chosen professional career; (*b*) having the dean of their school invite them to participate; and (*c*) telling the subjects that prior research had shown that skilled members of their profession are successful deceivers in this experiment (pilot studies had suggested that this was so). The balance of roles was achieved by instructing the subjects that part of their task was to detect whether the interviewer was being misled.

The deception was required to include both the withholding of an aroused negative affect and the simulation of another affect. If only the latter were required, there might be deception clues to the fact that the positive affect was simulated, and some leakage of anxiety or guilt about success in deception; but with the requirement that another affect be aroused and concealed, there would be greater likelihood of obtaining demonstrable leakage. A stress film was used to arouse negative affect, and the instructions required that the subject convince the interviewer that the subject was seeing a pleasant rather than a stressful film. A simple scenario was provided for the subject to use in the fabrication during the deceptive interview, to minimize differences which might occur as a function of subject's varying facility in spontaneously imagining what to say during deception. In the instructions requiring that the subject conceal negative affect and simulate positive affect, a brief description was provided of the content of the film they were to claim they had seen; a few positive emotions they were to claim they were experiencing were also described.

A less deceptive, more honest interaction was also arranged. A pleasant film was used to arouse positive affect, and instructions required that the subjects truthfully describe their reactions to the film. This was not, however, a totally relaxed interaction, since subjects believed that their ability to be convincingly truthful was under scrutiny and germane to their future career. And this more honest interaction was not expected to be totally free of deception, since subjects might well attempt to conceal any signs of anxiety about their skill in being convincingly truthful.

Four brief sessions, each lasting about seven minutes, were conducted consecutively within a one-hour period. In all four sessions, the subject first watched a film with the interviewer turned away from the subject and the film. After 60 seconds the interviewer turned to face the subject and continued to ask the subject questions while the subject viewed the film. At the end of the second minute, the film ended and the interviewer continued to ask questions. The interviewer was seated so that the film being viewed by the subject was never visible. In Session 1 and Session 2, the subject saw pleasant films and was instructed to describe any feelings frankly. In Session 3 and Session 4, the subject saw stressful films. In Session 3, the subject was instructed to conceal negative affect and to follow the scenario provided in convincing the interviewer that the subject was experiencing positive affect about a pleasant film. In Session 4, the subject was instructed to describe honestly any feelings about the stress film.

The instructions as to whether to be honest or deceptive were projected on the movie screen during the film viewing between the first and second minute of the film. After each session the subject answered an extensive written questionnaire about feelings experienced, feelings communicated, and the impression conveyed to the interviewer. In the introduction to the experiment, after the subject was given motivational information, the interviewer and subject were falsely told that some subjects would be asked to lie in all four sessions, some in no sessions, and some in just one or two sessions. They were also told falsely that the order of the sessions, both in terms of the film shown and the instruction provided, would vary across subjects. The purpose of this false information was to prevent the interviewer from knowing ahead of time whether a subject might be attempting deception in a particular session. As part of this attempt to mislead the interviewer, the actual content of the pleasant films shown in Sessions 1 and 2 was actually varied across subjects, as was the content of the scenario provided in the deception session. The schedule of questions followed by the interviewer was identical for all four sessions with two exceptions. In Session 3 (the deception session) and Session 4, the interviewer was instructed to challenge the subject by asking if the subject was really telling the truth, and also by asking if the subject would like to show the film to a young child. In the other two sessions, in place of these two questions the interviewer asked for associations to the film and descriptions of other comparative experiences.

The same adult female was the interviewer for all of the subjects across all of the sessions. Although identified to the subjects as a research social worker, she was actually a housewife with a bachelor's degree. She did not discover the actual order of the four sessions until she had seen two thirds of the subjects. Analysis of the judgments of observers to be reported below showed no difference between the subjects seen before and after the interviewer had discovered the order of the four sessions. The interviewer was naive and remained so about the hypothesis under study regarding the difference in the information conveyed by the face and the body.

Recordings

Two cameras, each connected to a separate videotape recorder, recorded the subject's behavior in the four sessions. One camera provided a closeup head-on view of just the subject's face. The other camera showed a head-on view of the entire face and body of the subject and a profile view of the face and body of the interviewer. These two recordings were synchronized by digital addresses placed simultaneously by a computer on both videotape recordings. The cameras were concealed. Although the subjects knew that an audio recording was being taken, the video recording was not mentioned until after the experiment.

Subjects

Beginning female nursing students were selected as subjects for three reasons. First, the ethical problems entailed in showing stress films could be reduced with this population. The stress films were composed of scenes from medical training films on amputations and the treatment of severe burns. Nursing students by virtue of their career choice have already elected to view such material at some point in their career. Second, a rationale could be readily provided for their engaging in deception; namely, a nurse must be able to deceive in certain situations. For example, when talking with the family of a severely injured child, she must conceal her own worries or distress and convey positive affect to reassure the parents. Third, the stakes for success in deception could be made high by telling the nurses that past research has found experienced nurses to do very well in the experiment and by implying that their behavior in the experiment was relevant to their future career.

The student nurses were invited by mail to participate in the experiment shortly after they had been notified of their acceptance in the School of Nursing, but prior to their matriculation. All 22 of the students living within commuting distance accepted. After the completion of the experiment, they were informed that their nonverbal behavior had been recorded by hidden cameras and that they had the right to ask that these records be erased. One of the 22 subjects chose to have her records destroyed. Five of the remaining 21 subjects confessed during the deception session. Their records were not considered in the data analysis to test Hypothesis 2, since by confessing they failed to provide an adequate sample of behavior while deception was in progress. Findings on the distinguishing characteristics of those who confessed are reported below.

Judgment Task

Two different judgment tasks were employed. In both, the observers were asked to judge whether a

sample of nonverbal behavior was from the more honest or the more deceptive interview; this was presented as a binary choice, honest or deceptive. In both tasks the same behavior samples were judged. In Task A the observers judged one sample of nonverbal behavior for each subject without any familiarity with the idiosyncracies in the subject's nonverbal repertoire. In Task B the observers first saw a videotaped example of the subject's behavior, which was identified as being nondeceptive, and then judged an unidentified sample.

No example of the subjects' usual behavior outside of this experiment had been videotaped, as would have been desirable for use as the example for familiarity. The closest approximation was the subjects' behavior during the two honest interviews, when they were shown a pleasant film and asked to describe their feelings honestly. Self-reports gathered after each interview rated the first honest interview as significantly less pleasant, relaxed, and natural than the second.[3] Both the example and the honest sample of the subject's behavior were therefore drawn from the second interview, in the first instance to approximate most closely the subject's usual behavior and in the second instance to provide maximum contrast with the deceptive interaction.

Drawing both the example and the unidentified sample from the same interview had undesired costs. Task B not only provided familiarity with the subject's behavior, it also increased the likelihood that the observers would base their judgments on the similarity or difference between the sample to be judged and the familiarity example, rather than on leakage or deception clues; that is, if a sample was similar, it was honest; if dissimilar, deceptive. While the use of any baseline sample would allow for such comparisons, the likelihood that the observers would employ such a strategy, and its chance of success, were heightened in this experiment. If the results on Task B were to show that the body was more accurately judged than the face, as predicted by Hypothesis 2, it would not be possible to ascertain whether this is so because the body provides more leakage and deception clues than the face, or more simply that between the deceptive and the honest interview, there are more differences in the appearance of the body than of the face.

The familiarity example was drawn from two episodes in the honest interview which the subject had rated as most pleasant: the subject's answer to the interviewer's first question during the film-watching portion of the interview ("What kinds of feelings are you having right now?") and the subject's reply to the interviewer's first question after the film was over ("How are you feeling now that the film is over?"). The unidentified sample from the honest interview was composed of two different episodes: the subject's response to the interviewer's second question during the film watching ("What kind of mood does the film create?") and to the second question after the film watching ("What other experiences have you had that conveyed the same feelings as the film?").

For the deceptive sample, two episodes were selected from the deceptive interview: the subject's response to the first question during the film watching ("What kinds of feelings are you having right now?") and to the question after the film watching ("Are you really telling me the truth? Do you think I believe you?").

The duration of the familiarity example was allowed to vary across subjects depending upon how long each subject took to reply, although it was always less than two minutes. The duration of the honest and deceptive samples varied across subjects but was matched for each subject by adjusting the length of the honest sample to that of the deceptive sample. All decisions about sampling were made from typescripts of the interviews without any inspection of the videotape records. Once sampling points had been determined, however, a few seconds of deviation were allowed, to avoid breaking an episode when the subject was moving.

Four videotapes were made for Task B, two showing the subject's face only and two showing the body only (the face was blocked off on the television monitor). Each tape contained samples of all 21 subjects, but a subject was shown only once (either the honest or deceptive sample) on a given videotape. Before seeing each sample to be judged, observers saw the familiarity example for that subject which was identified as an example of the subject's behavior when she was honest.

Six videotapes were made for Task A, three showing the face only and three showing the body only. In addition to the honest and deceptive samples, the familiarity examples were also presented for judgment. Again, each videotape contained the behavior of all 21 subjects, although only one sample (familiarity, honest, or deceptive) for each subject was contained in any one tape.

The instructions given to the observers provided as much information as has been given here about the nature of the honest and deceptive interviews. Observers were told that the purpose of the study was to examine what people could learn from observing the face or body, but the hypotheses were not mentioned. They were told that the length of the samples would vary, that one purpose of the experiment was to study how judgments are affected by the size of the sample, and that the duration of the sample was unrelated to whether the sample was honest or deceptive. The last was indeed the case. In order to decrease the likelihood that observers would attempt to balance their honest and deceptive judgments, they were falsely told that a tape might contain all deceptive samples, or all honest samples, or some of each. They were

[3] The differences on the ratings of the two honest interviews where a pleasant film had been shown were predicted. Pilot studies had suggested that in the initial session subjects were ill at east about what was required of them, and a second honest session was included in the experiment in order to obtain a more relaxed sample.

TABLE 1

SUMMARY TABLE FOR ANALYSIS OF VARIANCE

Source	df	MS	F
Between subjects	15	253.6	
Within subjects	112		
Area—Face or Body (A)	1	1194.4	9.28*
A × Subject	15	128.7	
Task—Task A or Task B (B)	1	453.8	12.77*
Condition—Honest or Deception (C)	1	2136.9	2.48
C × Subjects	15	863.3	
A × B	1	908.4	3.87
A × B × Subjects	15	234.9	
A × C	1	64.7	.12
A × C × Subjects	15	536.4	
B × C	1	21.9	.08
B × C × Subjects	15	265.9	
A × B × C	1	279.1	.93
A × B × C × Subjects	15	299.6	

*$p < .01$.

encouraged to judge each sample without regard for previous judgments.

Observers

The judgments were obtained during the usual meetings of upper-division psychology and speech classes at San Francisco State College. Observers were promised feedback about the findings of the experiment and their own individual accuracy scores to motivate their voluntary participation. In order to reduce the influence of any bias from any one class, the observers within each class were randomly assigned to two subgroups, one of which viewed a face tape, the other a body tape. After completing the judgment task, the observers were asked to rate the task on three 5-point scales: reasonable–unreasonable use of class time, boring–interesting, and frustrating–nonfrustrating. To reduce the likelihood of obtaining any purposefully random data due to annoyance with the task, the judgments of observers who said the task was extremely boring and unreasonable were discarded. Those discarded represented less than 5% of the total group of observers. In order to decrease the likelihood of obtaining random data because the observer did not understand the task or employed a different set of cultural rules in interpreting nonverbal behavior, the data from those who were foreign born were also discarded. In Task B, 55 observers remained who had judged the body tapes (28 who saw one body tape, 27 the other), and 48 remained who had judged the face tapes (25 who saw one face tape, 23 the other). In Task A, 65 observers judged the body tapes (24 who saw the first body tape, 22 the second, and 19 the third), and 65 observers judged the face tapes (20 who saw the first face tape, 19 the second, and 26 the third).

TABLE 2

CELL MEANS FOR ACCURACY OF JUDGMENT IN TERMS OF JUDGMENT TASK, TYPE OF NONVERBAL BEHAVIOR OBSERVED, AND INTERVIEW CONDITION JUDGED

Judgment Task	Interview condition	Type of behavior observed	
		Face	Body
Task A (no familiarity sample)	Honest	42.50	44.81
	Deceptive	51.37	50.63
Task B (familiarity sample)	Honest	43.06	50.13
	Deceptive	47.69	63.50

RESULTS

The data for testing Hypothesis 1 came from the subjects' responses to one of a series of questions asked immediately after the end of the experiment. Each subject was asked to list what she thought one should control or avoid doing to be a successful deceiver. As predicted, more mentioned the face (17) than mentioned the body (4) ($\chi^2 = 12$, $p < .01$).

The data for testing Hypothesis 2 came from the observers' judgments of the videotapes. The judgments within each group of observers were tallied to provide the proportion of correct judgments for each person they had viewed. Four such accuracy scores for each subject were derived from the observers who had utilized Task B: one when the person's face during the honest sample was judged (face honest), one for her face during the deception sample (face deceptive), and one each for body honest and body deceptive. Four more such accuracy scores were derived from the observers who had utilized Task A: face honest, face deceptive, body honest, body deceptive. These eight scores on each of the 16 subjects who had not confessed were entered into a $2 \times 2 \times 2$ repeated-measurements analysis of variance, in which the variables were area (face or body), task (A or B, i.e., with or without a familiarity example), and condition (whether the behavior judged was actually honest or deceptive). As noted earlier, the judgments of the five subjects who confessed were not included in this data analysis.

Table 1 shows that there was a significant main effect for area and for judgment task. Inspection of the cell means in Table 2 and the t tests between means showed, however,

that these main effects were not general across experimental conditions, but were due to the relatively high accuracy achieved in one cell—when the body during deception was judged by observers using Task B. There was no significant difference in the accuracy achieved from the face or body on Task A. On Task B, which employed the familiarity example, the accuracy achieved from the body was significantly greater than from the face ($t = 2.60$, $p = .01$) when the deceptive behavior was judged. There was a similar trend when the honest behavior was judged, but it did not reach significance ($t = 1.01$, $p = .20$).

These results provided partial support for Hypothesis 2, with two limitations. First, the predicted face–body difference was not common across both judgment tasks, but limited to Task B, which employed the familiarity example. Second, the face–body difference was not restricted, as had been predicted, to judgments of deceptive behavior. There was a similar trend for judgments of the honest behavior on Task B.

The main effect for judgment task in the analysis of variance was not general across experimental conditions but like the results for area, was due to accuracy in one condition. It was only when the body behavior during deception was judged that the results for Task B were significantly different from the results for Task A ($t = 2.29$, $p < .05$).

Failure to maintain deception had not been expected for many of the subjects, but 5 of the 21 subjects did confess that they were lying. The results of observers' judgments of the confessors' face and body during the honest and deceptive interviews were excluded from the data analysis testing Hypothesis 2 because these subjects confessed so early in the deceptive session that only a very brief sample of their behavior prior to confession was available. While this experiment was not designed to investigate the characteristics of those who confessed, it was of interest to learn about the segment of the population for whom Hypothesis 2 does not apply. One possibility was that the confessors were more upset by the stress film than were the nonconfessors and failed to maintain the deception because of greater negative affect arousal. The data available failed to support this explanation. After each interview the subjects had rated their feelings about the films on 10 unipolar emotion scales. There were no significant differences between the confessors and nonconfessors.

There were no other data available on the subjects. Despite the inherent problems in ex post facto testing, all 21 subjects were contacted six months later to explore individual differences. A number of tests were included which colleagues had suggested might explain why some of the subjects had confessed. The results were notable for a lack of differences between confessors and nonconfessors. They did not differ on the Vocabulary or Comprehension test from the Wechsler Adult Intelligence Scale (Wechsler, 1955), the Mach V (Christie, 1970), the Interpersonal Check List (Leary, 1955), the Gottschaldt Figures Test (Crutchfield, 1950), Thurstone's adaptation of the Stroop Color-Word Test (Thurstone, 1944), or on 15 of the 18 California Psychological Inventory scales (Gough, 1957). The confessors had higher scores than the nonconfessors on the Self Acceptance scale ($t = 2.91$, $p = .01$), the Achievement via Independence scale ($t = 3.87$, $p = .01$), and the Flexibility scale ($t = 2.51$, $p = .05$). Because these differences on the California Psychological Inventory had not been predicted, because most of the scales on this test failed to show differences, and because the other personality, cognitive-style, and intelligence tests also failed to show differences, it is hard to place much confidence in these few significant findings.

Another possibility explored was that the confessors might have been less motivated to become nurses, or to participate in the experiment, and therefore did not work as hard in the most difficult phase of the experiment, maintaining deception. Eight questions designed to measure motivation to become a nurse failed to show any difference. Two questions about the relevance of the experiment to nursing did show differences. The confessors showed significantly less agreement with the statements, "The ability to conceal feelings in this experiment is relevant to evaluating most people's potential as a nurse"

($t = -2.57$, $p = .05$) and "to evaluating my potential as a nurse" ($t = -4.34$, $p = .01$). On the other hand, the confessors, like the nonconfessors, strongly agreed with the statement, "I thought my participation in this experiment was very relevant to my professional goal of becoming a nurse." Also the confessors, like the nonconfessors, strongly endorsed the statements, "The ability to conceal is very relevant in evaluating the potential of a nurse" and "In my intended area of nursing specialty it is very important to be able to conceal one's feelings." It would seem then that the confessors were similar to the nonconfessors in motivation to become nurses, in believing the ability to conceal is relevant to nursing, and in thinking the experiment was useful. But the confessors did not think their failure to maintain deception in the experiment was a mark against them. This finding seems particularly vulnerable to after-the-fact testing. The confessors, more than the nonconfessors, reported, "I tended to worry about the experiment a lot after it was over" ($t = 3.64$, $p = .01$).

In interviews immediately after the experiment, most of the confessors described themselves as generally unable to hide their feelings from others. On the basis of this lead, a true–false questionnaire was composed of eight items, such as "If I want to I can fool others and make them think I feel a certain way when I really don't." On this test of self-reported ability to control the display of feelings, the confessors received a lower score than the nonconfessors ($t = -5.45$, $p = .01$), and there was no overlap between the scores of the confessors and nonconfessors. This difference was supported by the finding that confessors agreed less than the nonconfessors with the statements, "In professional life I conceal very well" ($t = -3.36$, $p = .01$) and "In personal life I conceal very well" ($t = -2.54$, $p = .05$).

Discussion

The main purpose of this experiment was to test hypotheses derived from Ekman and Friesen's (1969) theory of nonverbal leakage about differences between the face and body. There are two central and logically related propositions in that formulation. The first proposition maintains that people are generally more aware of their facial behavior than their bodily activity and therefore will be more likely to disguise the face than the body. The results confirming the first hypothesis supported this thinking. Among subjects who had attempted to deceive, more mentioned attempting to disguise the face than the body.

The proposition that people attempt to lie more with the face than the body is the basis for the second proposition, that the body more than the face gives deception clues and leaks the withheld information. On the basis of such proposed differences in facial and body behavior, the second hypothesis predicted that observers' judgments of whether someone is honest or deceptive would be more accurate if deceptive behavior was judged when the body was viewed than when the face was viewed, but not if honest behavior was judged. The results only partially supported the second hypothesis. There was no difference in accuracy between the body and the face when the observers had no prior familiarity with the subjects' nonverbal repertoire. It was only when the observers saw a sample of each person's behavior from the honest interview which was identified as nondeceptive behavior that more accurate judgments were made from the body than from the face. In that judgment task the observers may well have made their decisions by comparing the behavior to be judged with the identified honest sample, calling the unidentified sample honest if it appeared similar and deceptive if it looked different. The fact that accuracy was higher for the body than for the face under such circumstances means that comparisons were of little utility for the face but profitable for the body. Presumably this difference is explained by the reasoning which led to Hypothesis 2. If people do indeed lie with the face, then comparing the face during deception with the honest face would not be of much use, since the facial behavior in deception would be disguised to look like the honest example. If people do not so manage their body behavior, then comparing a sample of body behavior during deception with an example of honest body behavior should reveal differences.

While judgments were more accurate from the body than the face, in accordance with Hypothesis 2, this was so only when Task B was employed, during which the observers could compare the behavior they judged with the familiarity example; therefore, ambiguity remains about the basis of the more accurate judgments. Is it that the observers recognized the signs of leakage and deception clues when they compared the body behavior during deception with the familiarity example? Or, is it more simply that the observers noticed differences between the body behavior shown in the deceptive interview and the familiarity example, without reading these differences as leakage or deception clues? While either possibility is consistent with the reasoning which led to Hypothesis 2, the data from the experiment do not allow a choice between these explanations.

Other studies of the honest and deceptive interviews will clarify this issue and more directly test Ekman and Friesen's theory of nonverbal leakage and deception clues. Measurements of the behavior itself should show similarity in the facial behavior between the honest and deceptive interviews and show difference in the body behavior. Such measurement is under way, but only hand movements have been analyzed to date. Results are encouraging, showing the theoretically predicted changes in particular classes of hand movements from honest to deceptive interviews.[4] Another approach would be to show observers samples of honest behavior and samples of deceptive behavior without providing any information about the situation. Judgments on scales such as relaxed-tense and pleasant unpleasant should show more similarity between honest and deceptive facial behavior than between honest and deceptive body behavior.[5]

The fact that there was no difference in accuracy between the face and body when the behavior was judged without the opportunity for comparison with an example of honest behavior suggests that the differences in body behavior during deception were in the form of deception clues rather than leakage. If there had been leakage of negative affect, presumably the observers would have known the subject was being deceptive without needing to compare the body with the honest example. If the body instead provided only deception clues—some sign that something was amiss but not the actual negative affect—then judging these body movements as evidence of deception might not have been possible without familiarity with the person's body behavior when honest.

The findings were somewhat equivocal in regard to the part of Hypothesis 2, which had predicted that accuracy would be greater for the body than for the face only when deceptive behavior was evaluated. Among observers who had been given a familiarity example, there was a significant difference between body and face, as had been predicted, when the deceptive behavior was judged; but contrary to expectation, there was a marked, though not significant, trend in the same direction when the honest behavior was judged. The nature of the judgment task may have worked against Hypothesis 2. Since the judgment choice was binary, deceptive or honest, if the body did provide accurate information when someone was most deceptive, observers might have made accurate decisions about the more honest body behavior by exclusion. They could have reasoned that since a given sample of body behavior did not show this or that sign of leakage, they should call it honest. Thus, accuracy in judging the honest behavior could be inflated because of the accurate information available in the deceptive body behavior.

Let us consider the results on the judgments of the face. Not only were the facial judgments less accurate than the body judg-

[4] Findings to date show that what Ekman and Friesen (1969) have defined as hand *illustrators* decreased from honest to deceptive interviews, while the hand shrug *emblem* increased and the face-play *self-adaptor* increased. The analysis of postural shifts, though not complete, suggests that there are more shifts in the deceptive than the honest interviews. The analysis of leg movements, head orientation, and facial expression will also be undertaken. When all of these measurements are complete, they will provide the most direct test of Ekman and Friesen's theory of nonverbal leakage and deception clues.

[5] Subsequent to the submission of this report, we have conducted this study and have obtained findings as predicted.

ments, but when the face was viewed, regardless of whether a familiarity example was seen or whether the behavior was from the honest or deceptive session, the number of persons whose behavior was accurately judged by more than half of the observers was no better than would be expected by chance.[6] Despite these poor results, two additional findings suggest that the facial behavior during deception was not totally devoid of meaning. First, other observers more highly trained were able to judge this facial behavior accurately. Ekman and Friesen's theory suggested that facial behavior during deception would contain leakage in microexpressions, very brief facial movements shortened by attempts to inhibit or interrupt, but that these would be missed by most observers. To check this possibility, four experienced facial analysts who had spent hundreds of hours using a procedure for measuring facial movements (Facial Affect Scoring Technique, Ekman, Friesen, & Tomkins, 1971) were shown the videotapes of the face with the familiarity example. Each of these observers accurately judged both the honest and deceptive behavior of almost all of the nurses, suggesting that the information is there in the face, if the viewer knows what to look for and how to interpret it. Obviously, this finding must be considered as very tentative until it has been replicated with more observers and more subjects.

The second finding which suggests that the facial behavior during deception was not devoid of meaning comes from correlations between how the face during deception was judged and other aspects of the subjects' lives. Inspection of the judgment data revealed large individual differences among the subjects. Some subjects evidently were good facial liars, since the great majority of the observers called them honest during the deception session. Others were poor facial liars, since the great majority of the observers correctly judged their behavior during deception. And some subjects' facial behavior must have been contradictory or confusing, since the observers were about evenly divided as to whether these subjects were honest or deceptive. These differences among subjects in how their facial behavior was judged, and presumably in how well they censored and simulated their facial expressions, were found to be correlated with a number of personality variables and subsequent performance measures. These results will be the subject of a later report, when the findings have been replicated with other subjects to provide a larger sample. Only one correlation with subsequent school performance is reported here, to illustrate that the individual differences in facial behavior during deception were not a chance phenomenon. This particular result is reported because it had been predicted, is straightforward, and perhaps best demonstrates that behavior in the honest–deception experiment was relevant to the rest of the subjects' lives.

It was reported in the Method section that the subjects had been told that behavior in the honest-deceptive session was relevant to success in nursing, that good nurses had to be able to reveal and conceal, and that experienced nurses did well in this task. At the time, such claims were based largely on conjecture, bolstered only tentatively by some pilot studies. The results now show that this is very likely the case. How the subject's facial behavior during deception was judged—whether the subject was called honest or deceptive—was correlated with academic and clinical grades in all three quarters of the following year. For example, the supervisors' ratings of the subject's work with patients one year later was positively correlated with the subject's being a successful facial deceiver; the correlation was .55 with the judgments of the facial behavior during deception when the observers had not seen a sample to provide familiarity, and .62 with the facial judgments when a familiarity sample had been seen.

To review this discussion of the judgments of facial behavior, there was tentative evidence to suggest that observers who have received special training can accurately detect deception from facial behavior and that naive observers cannot. Persons differed in how

[6] To answer this question in another fashion, the data were also tallied by observer rather than by subject, counting how many subjects each observer accurately judged, rather than counting how many observers were accurate in judging each subject. Identical results were found.

their facial behavior during deception was judged by untrained observers, and these judgments were related to measures of subsequent competency in nursing. Skill in facial deception (observers mistakenly judging the subject's face during deception as honest) was correlated with ratings of working well with patients one year later. This finding suggests that how the subjects behaved in this experiment, as measured from their nonverbal behavior, was not irrelevant to the rest of their life, but instead was predictive of their subsequent career.

Five of the 21 subjects confessed during the deceptive session, and thus were not included in the data for testing Hypothesis 2. Comparisons of these subjects with the majority who did not confess suggested that confession could not be attributed to being more upset by the stress film shown in the deceptive session or by the difference in motivation to become a nurse. A number of other possibilities were explored with meager results, except that confessors, more than nonconfessors, reported being typically unable to hide their feelings from others. Further exploration of the characteristics of those who cannot maintain deception, for whom the face-body hypothesis is not applicable, will have to wait replication of the experiment with more subjects so that a larger group of confessors can be studied and other possible characteristics investigated.

In conclusion, the proposition that the face more than the body is subject to control and disguise during deception was partially, indirectly supported. When asked what behavior they should control when deceiving, subjects mentioned the face more often than the body. When deceptive behavior was judged and observers had seen an example of each subject's honest behavior which was so identified, judgments of the body were more accurate than of the face. Comparing facial behavior during deception with an example of honest facial behavior did not yield accurate judgments, and the judgments were no more accurate than those made without any such knowledge of the person's honest facial behavior, presumably because the subjects disguised their facial behavior during deception to appear honest. Comparing body behavior during deception with an example of honest body behavior did yield accurate judgments, and the judgments were more accurate than those made without any such knowledge of the person's honest body behavior, presumably because the subjects did not disguise their body behavior during deception, and it therefore looked different from their honest body behavior. Although the evidence is limited because of the small number (16) of persons judged, it is noteworthy that similar findings were obtained in a past study with different subjects in a different type of deceptive situation. In an earlier study (Ekman & Friesen, 1969) the body was also found to provide more accurate information than the face when observers judged naturally occurring deception shown by three female psychiatric patients during psychiatric interviews.

The findings from this experimental deception situation need to be replicated with additional subjects and observers. Furthermore, the underlying proposition about the difference in disguising facial and body behavior during deception which was the basis for Hypothesis 1 (subjects will more often report that they disguise the face during deception than that they disguise the body), and Hypothesis 2 (observers' judgments of deception will be more accurate from the face than from the body) needs to be directly approached. Studies actually measuring the occurrence of particular body movements and facial expressions during the honest and deceptive interviews are in progress.

REFERENCES

CHRISTIE, R. Scale construction. In R. Christie & F. L. Geis (Eds.), *Studies in Machiavellianism*. New York: Academic Press, 1970.

CRUTCHFIELD, R. S. *Gottschaldt Figures Test*. Berkeley: University of California, Institute of Personality Assessment and Research, 1950.

DARWIN, C. *The expression of the emotions in man and animals*. New York: Philosophical Library, 1955. (Originally published: 1872.)

DUNCAN, S. Nonverbal communication. *Psychological Bulletin*, 1969, **72**, 118–137.

EKMAN, P., & FRIESEN, W. V. Nonverbal leakage and clues to deception. *Psychiatry*, 1969, **32**, 88–106.

EKMAN, P., FRIESEN, W. V., & ELLSWORTH, P. *Emotion in the human face: Guidelines for research and an integration of findings*. New York: Pergamon Press, 1972.

EKMAN, P., FRIESEN, W. V., & TOMKINS, S. S. Facial Affect Scoring Technique (FAST): A first validity study. *Semiotica*, 1971, 3, 37–58.

FAST, J. *Body language*. New York: Evans, 1970.

FELDMAN, S. S. *Mannerisms of speech and gestures in everyday life*. New York: International Universities Press, 1959.

FREUD, S. Fragment of an analysis of a case of hysteria. In, *Collected papers*. Vol. 3. New York: Basic Books, 1959.

GOFFMAN, E. *The presentation of self in everyday life*. New York: Doubleday, 1959.

GOUGH, H. H. *Manual for the California Psychological Inventory*. Palo Alto, Calif.: Consulting Psychologist Press, 1957.

HARRISON, R. Nonverbal communication. In I. de Sola Pool et al. (Eds.), *The communications handbook*. Chicago: Rand McNally, 1973.

KNAPP, M. L. *Nonverbal communication in human interaction*. New York: Holt, Rinehart & Winston, 1972.

LEARY, T. Interpersonal Check List. Berkeley, Calif.: Psychological Consultation Service, 1955.

RUESCH, J., & KEES, W. *Nonverbal communication: Notes on the visual perception of human relations*. Berkeley: University of California Press, 1956.

THURSTONE, L. L. *A factorial study of perception*. Chicago: University of Chicago Press, 1944.

WECHSLER, D. *Wechsler Adult Intelligence Scale: Manual*. New York: Psychological Corporation, 1955.

(Received June 30, 1972)