

BODY POSITION, FACIAL EXPRESSION, AND VERBAL BEHAVIOR DURING INTERVIEWS¹

PAUL EKMAN²

School of Medicine, University of California, San Francisco

The communicative value of body position and facial expression was evaluated by measuring an O's ability to detect a relationship between nonverbal and verbal behavior which had been simultaneously emitted. The verbal and nonverbal stimuli were collected during 2 different standardized stress interviews. Judges (Js) were shown pairs of photographs together with short written speech samples and required on each trial to pick the photograph which matched the verbal behavior. In 4 separate experiments with different groups of Js, accurate judgments were obtained. Evidence for a relationship between nonverbal and verbal behavior simultaneously emitted was replicated across 2 different samples of interview behavior and under 3 cue conditions—seeing the head, body, or whole person.

There has been relatively little systematic investigation of the information which may be transmitted through spontaneous nonverbal behavior shown during interpersonal transactions. Research on body movement and facial expression has had to deal with a phenomenon which is continuously occurring, has no readily apparent unit of measurement or method of evaluation, and is both difficult and expensive to record. Despite the clinical conviction that nonverbal behavior provides important clues about an individual's emotional experience, such beliefs rest on little more than anecdotal evidence or speculation. In fact, research on interview behavior has increasingly focused on verbal rather than nonverbal interactions.

The major problem in exploration of the nonverbal aspects of interview behavior, may at least initially appear to be the acquisition of a permanent record. Actually, unless the body movements and facial expressions are obviously meaningful from simple inspection, the investigator is quickly overwhelmed with a mass of photographic stimuli, which are just as baffling and complex as were the original behaviors themselves. The central dif-

ficulty in research on spontaneous nonverbal behavior is the development of fruitful techniques for evaluating the information which may be contained in records of nonverbal behavior. Two approaches can be distinguished which although related, usually have been separately pursued.

The first approach has measured variations in nonverbal behavior as a *response* to some other factor, such as interview structure or patient mood. Sainesbury (1955) and Dittmann (1962), for example, have reported a relationship between changes in nonverbal behavior and the content or structure of an interview. They did not study the related question, however, of what may be communicated to an observer by these changes in nonverbal behavior.

In the second research approach, nonverbal behavior is presented as a communicative *stimulus* and an observer's response to that stimulus is measured. Most of the experiments using this second approach have been performed with posed behavior which occurred during isolation rather than in the midst of an interpersonal interaction (for example, Schlosberg, 1954). Moreover, these studies have focused upon judge agreement rather than accuracy in understanding communication through nonverbal behavior. Attempts to measure the information transmitted through nonverbal behavior during an interview have been inconsistent in their methods and results (Giedt, 1955; Mahl, Danet, & Norton, 1959). These experiments

¹ This project was supported by a postdoctoral research fellowship, MF 6092, from the National Institute of Mental Health, United States Public Health Service. The author is indebted to Sanford Autumn for his many valuable suggestions about the research design, statistical analysis, and preparation of the manuscript.

² Now at San Francisco State College.

have not allowed specification of the actual nonverbal cues which served as the basis for the judgments required. A further criticism of the research studying nonverbal behavior as a communicative stimulus, is that it has relied on the application of highly complex verbal concepts for evaluating the information transmitted by the nonverbal cues. A judgment task which requires ratings on emotional scales, diagnostic or psychodynamic formulations, may well miss much of the information which nonverbal behavior can communicate to an observer.

The present series of experiments combined features of both research approaches by measuring an observer's response to nonverbal behavior (Approach 2), in a task which required him to relate nonverbal behavior to another aspect of the interview situation (Approach 1). A relationship between what a person does with his body and what he is saying was assumed, and a matching procedure was borrowed from the early research on expressive behavior (Vernon, 1936; Wolff, 1943) to test the following hypothesis: Judges can detect a relationship between nonverbal and verbal behavior which was simultaneously emitted during an interview.

Four separate experiments with different groups of judges will be reported. In each of these experiments judges responded to verbal and nonverbal stimuli which had been gathered during the course of two interviews.

GENERAL METHOD

Interview Procedure

A standard rather than clinical interview was employed to increase the comparability of the two interviews and to insure the presence of different affective reactions. The interview structure was achieved by programing both the style and content of the interviewer's (the experimenter) behavior toward the interviewee (the subject). After an introductory affectively neutral period of 10 minutes, the experimenter attacked and criticized the subject's choice of occupation, competency, and motivation. Throughout this stress phase, the experimenter continually questioned and interrupted the subject, responding tangentially to his replies, and generally giving little opportunity for any defense. After 10 such minutes, the experimenter initiated the catharsis phase by apologizing for being hostile, explaining it had been necessary in order to study the subject's reaction to stress. The subject was praised for his resiliency and performance under stress, and humor

was introduced. Throughout the following 10 minutes of catharsis, the experimenter was reassuring in manner and attempted to bring about a release of tension.

Interview Participants

Two such standardized interviews were recorded, with completely different participants in each of the interviews. All four interview participants were from the field of psychology. The two different experimenters were staff research psychologists at a veteran's hospital, having a position of some authority over the two different trainees who served as subjects. All of the participants were told that the experiment was concerned with interviewing, and that their behavior would be tape recorded and observed. The experimenters also knew that they would be photographed and nonverbal behavior would be studied.

Recording Method

Photographs showing a profile view of both the experimenter and the subject were taken through a one-way vision screen with a 35-millimeter still camera. The verbal behavior was tape recorded. The verbal and nonverbal records were synchronized by a switching device mounted on the camera which placed an audible signal over the tape recorded voices whenever a picture was taken. The frequency of photographic sampling, one frame every 30 seconds in Interview A and one frame every 15 seconds in Interview B, was dictated by limitations in the film capacity of the camera.

Selection of Verbal and Nonverbal Stimuli

The judgment task involved presenting short written speech samples from the interviews along with pairs of photographs. On each trial the judge was required to pick from a pair of pictures the one which best fitted or matched the verbal behavior.

The universe of verbal behavior which could be sampled for presentation in the judging task was delimited by the number of occasions during the interview in which a photograph had been taken. There were 26 such points during Interview A and 71 points during Interview B. Speech samples of from 30 to 60 seconds were chosen from this pool of verbal behavior if they appeared intelligible when lifted from the context of the total conversation. There was no inspection of the photographs during this selection procedure. An index card was prepared for each speech sample which stated whether the verbal interaction had been taken from the stress or catharsis phase of the interview; provided a one-sentence background synopsis; contained the speech sample of one or two statements by both the experimenter and the subject, with the actual moment which had been photographed indicated by a circled word.

Pairs of photographs, correct and incorrect, were then selected to accompany each speech sample.

The correct photographs were the ones which actually had been taken during the speech samples. Incorrect photographs were randomly selected from a photograph pool delimited by two factors: The mouth position, open or closed, had to be the same in both the incorrect and correct photographs, and if the experimenter (or the subject) were speaking when the correct photograph was taken, then the incorrect photograph had to be selected from photographs also taken while the experimenter (or the subject) was speaking.

Two types of incorrect photographs could be chosen for pairing with each correct one. The incorrect photograph could be taken from the same phase of the interview as the correct picture, or taken from the other phase. Since speech samples had been selected from both the stress and catharsis phases of the interview, there were four types of correct-incorrect photograph pairs which could be presented to the judges. With a verbal sample from the stress phase of the interview, the judges could be required to choose between two stress photographs or between a stress and a catharsis pair. Similarly, with a verbal sample from the catharsis phase of the interview, the judges could be presented with either two catharsis photographs or with a catharsis and a stress photograph.

Experiment I

The problem was to determine if any relationship between verbal and nonverbal behavior simultaneously emitted is communicated to an observer. Hypothesis: untrained judges can choose from a pair of photographs the one which was taken during a given speech sample.

Subjects. Eighteen college freshman, consisting of the entire introductory psychology class taught by the author, served as judges. There were 10 males, 8 females; the median age was 18.5.

Method. Fourteen speech samples and accompanying correct-incorrect photographs were selected from Interview B by the method described above. Eight of these speech samples were from the stress phase of the interview; six, from the catharsis phase. Four × five inch enlargements were made of each photograph, showing a profile view of both the experimenter and the subject. The judgment task was individually administered in order to allow random variation of the order of presenting the 14 speech-photograph items. The judge was shown pairs of photographs, each of which came from the same interview phase on half the items, while on the remaining trials the incorrect-correct photograph pairs were drawn from different interview phases. A table of random numbers was used to determine whether for a particular judge the choice on a particular speech would be between photograph pairs from the same or different interview phases.

Prior to the task, the judge was given the following instructions to read:

This is a test of your skill in interpreting and understanding gestures and body movements. You

will be shown some photographs which were taken every 15 seconds during a thirty minute interview. The interviewer, or Examiner, was a staff psychologist at a hospital, and the person interviewed, the Subject, was a student in training to become a psychologist. The Subject was told that he was participating in a research project on interviewing techniques, and that he would be observed through a one-way vision screen and tape recorded. The interview was pre-arranged for the Examiner to ask factual questions during the introductory first 10 minutes, and then to become hostile and challenging, questioning the Subject's academic training, the adequacy of his preparation to take his examinations coming up the following month, and to continually interrupt him. This Stressful period lasted about 10 minutes. In the Final phase of the interview the Examiner explained what he had been doing; that it had been part of the research to try and provoke the Subject, and generally attempted to reassure the Subject. In actuality, this interview plan did not work out perfectly. There was not as clearly defined a difference between the three phases of the interview as had been expected. There was some stress for the Subject throughout the interview. The final phase was not completely successful in producing some relief from the stress since the Subject knew that the experiment was still continuing, and perhaps was apprehensive about what might be coming next. Nevertheless, there were some important differences between the various phases of the interview. The stress phase did have more overt expression of hostility and tension; and in the final phase the Subject did experience some relief, at least knowing that the worst was over. You will be given to read 14 excerpts from the tape-recorded conversation. The verbal excerpts each took less than a minute of interview time. One photograph was taken during each of the verbal excerpts, and the actual words spoken when the photograph was taken are circled in red. After you read each verbal excerpt you will be shown a pair of photographs, and your job as a judge will be to pick which of the two pictures in the pair fits the best or matches the verbal excerpt. In all of the photographs, the Examiner appears on the left and the subject on the right. Your task, then, is to determine for each set of two photographs which one is most likely to have been taken during the verbal excerpt.

Results. The number of correct or accurate choices for each judge was tabulated. This distribution of obtained accuracy scores was evaluated by comparison with a theoretical median representing the score which might be expected if only chance factors were operative. An assumption was made that if there had been no information in the situation, and the judges' choice had simply re-

TABLE 1
RESULTS OF THE FOUR EXPERIMENTS

Trials	Experiment I	Experiment II		Experiment III		Experiment IV	
	N = 18	N = 16		N = 15		N = 27	N = 29
	Whole person	Body	Whole	Body	Whole	Body	Head
Interview A (6 trials) Obtained median % above expected median of 3		3.2 44	4.0* 56	2.5 13	3.6 56	2.6** 7	3.9** 76
Interview B (14 trials) Obtained median % above expected median of 7	9.5*** 78	7.5 50	8.2*** 69	9.1*** 75	8.1 56	8.1* 67	7.7 55
Interviews A and B combined (20 trials) Obtained median % above expected median of 10		10.7* 62	12.7* 75	11.2*** 67	11.8*** 60	10.6 52	11.3* 73

* $p < .05$.
** $p < .01$.
*** $p < .001$.

flected chance factors, then there would have been an equal probability of either photograph being selected on each trial. The expected median was therefore equal to a correct choice on half of the trials. Significance tests were derived by applying Wilcoxon's matched-pair ranks test (Siegel, 1956, p. 75) to the difference between the obtained scores and the expected median.

It can be seen in Table 1 that the obtained median accuracy score was significantly better than might be expected by chance. Table 1 also shows that more than three-fourths of the judges scored above the expected median.

Experiment II

Problem. The purpose of this experiment was to replicate the findings reported above with a different group of judges and to extend the generality of the results by utilizing verbal and nonverbal stimuli drawn from an additional interview. A second purpose was to determine if it were still possible to match verbal and nonverbal behavior with more restricted nonverbal cue information. In a pilot study (Ekman, 1961) an attempt was made to specify the nonverbal cues required for the judgment task by limiting the judges to seeing either head or body. The results indicated little difference in accuracy if the judge saw head or whole person cues, but accuracy was much lower with photographs showing only body position. Hypothesis: Judges can choose from a pair of photographs the one which was taken during a given speech sample when responding to pictures of the whole person. This discrimination between photograph pairs is

also possible when the judges are limited to pictures showing only body position.

Subjects. Since the pilot study had suggested that working with body position photographs was a difficult task, an attempt was made to select as judges individuals who might be able to best understand this type of nonverbal behavior. Sixteen professional modern dancers served as judges; 14 were female, the median age was 25, the median education was 16 years.

Method. The same task was employed, but the procedure was modified for group presentation rather than individual administration. A booklet was prepared which contained an instruction sheet, similar to the one used in the first experiment, and 20 speech samples. Fourteen of these were the samples used in the first experiment, the other 6 speech samples were drawn from Interview A. Twelve of the 20 samples were from the stress phases of the interviews, the remaining 8 samples were drawn from the catharsis phases. Half of the correct-incorrect photograph pairs were composed of pictures taken during the same interview phase, the remaining 10 pairs of photographs consisted of pictures from different interview phases. The photographs were made into two sets of positive transparencies. The heads of both the experimenter and the subject were covered with opaque ink in one set of transparencies, so that only the body position could be seen.

Two projection screens and two 35-millimeter slide projectors were used to show the correct-incorrect photograph pairs simultaneously to the group of judges after they read each of the speech samples in the booklet. The judges first completed the 20 trials responding to photographs of body position, and then after a 10-minute intervening task, repeated the procedure with photographs showing the whole person. The appearance of the

correct photograph on either the left or right projection screen differed under the body and whole person cue conditions and was balanced across the 20 trials.

Results. The technique of data analysis was the same as described in the first experiment. The results on Interview A trials and Interview B trials in this experiment and in Experiments III and IV will be interpreted together in the discussion section of this paper. The results in Table 1 show that accuracy on all 20 trials (Interviews A and B combined) replicated the findings of the first experiment. The judges were able to pick the photograph which matched a verbal sample when limited to seeing body position and also when shown the whole person. While it appears that the level of accuracy was greater when the whole person was seen, the difference between the two cue conditions was not tested since there had been no control for practice effects.

Experiment III

Problem. In interpreting the results from Experiment II, it was not clear whether the judges' ability to match "body only" photographs with speech samples was due to any special characteristics which might be associated with dancers or if such accuracy might be found in a more unselected group of judges. The major purpose of this experiment was, then, to determine if a heterogeneous group of judges could match body position photographs with speech samples. A secondary purpose was to replicate the findings of the two earlier experiments. Hypothesis: A heterogeneous group of judges can choose from a pair of photographs, showing only the body position, the one which was taken during a given speech sample.

Subjects. Fifteen undergraduate students in the author's introductory psychology evening class served as judges. Twelve were males; the median age was 22.5 with a range from 18 to 55.

Method. Exactly the same procedure described in Experiment II was followed. Twenty speech samples were presented in booklet form. The judges first responded to body position photographs, and then to pictures showing the whole person.

Results. The techniques of data analysis were identical to those reported earlier. Table 1 shows that for the results of all 20 trials, the hypothesis was confirmed and the earlier findings were replicated. No comparison was made between the accuracy achieved by the dancers in Experiment II and this group of evening psychology students, since the two

groups of judges differed in age, education, and sex distribution as well as in their choice of vocation.

Experiment IV

Problem. The design in Experiments II and III did not permit study of relative accuracy in matching nonverbal and verbal behavior as a function of the nonverbal cue information available. The purpose of this experiment was to test impressions from the pilot study referred to earlier by comparing accuracy between groups of judges who were limited to seeing either head or body position nonverbal cues. Hypothesis: Judges can choose from a pair of photographs the one which was taken during a given speech sample when responding to either head or body position cues. The judges shown head cues will achieve greater accuracy than the group of judges seeing body position cues.

Subjects. Female students in two freshman psychology classes served as judges. There were 27 judges in the group seeing body position photographs, 29 judges in the group responding to head photographs. The median age in both groups was 18.

Method. The general procedure described in Experiments II and III was followed except that two separate groups served as judges, each group responding to only one cue condition, either head or body.

Results. The technique of data analysis described earlier was again used to evaluate the results which are shown in Table 1. Considering the group which responded to head cues, the results on all 20 trials combined generally conform to the findings reported earlier for judgments of the whole person. The group responding to body position cues, however, did not achieve a significant level of accuracy in their scores on all 20 trials. Examination of the results on Interviews A and B separately revealed that while accurate judgments were achieved on Interview B, the judges were significantly (two-tailed test) below the chance level of expectation in their responses to body position photographs from Interview A. There was a tendency for body photographs from Interview A to be similarly misjudged in Experiment III, although it had not reached statistical significance.

Differences in accuracy under the two cue conditions were evaluated with a Mann-Whitney *U* test. The group responding to head cues achieved greater accuracy than the body position judges in their performance on all 20 trials ($p < .05$) and in their judgment of photographs from Interview A ($p < .001$).

DISCUSSION

Accuracy in matching materials from the two different interviews was related to the type of cue information provided—head, body, or whole person cues. Although there was some inconsistency across the four experiments, by and large a significant level of accuracy was achieved on both Interviews A and B if either the whole person or the head were shown. There were marked differences in the accuracy obtained on Interviews A and B, however, when only the body positions were shown. The body cues in Interview B seem to have provided sufficient information for matching in two out of three experiments. Table 1 shows that such accuracy was never reached on body cue judgments of Interview A. This difference in accuracy was evaluated with the McNemar test for the significance of change (Siegel, 1956, p. 63) by comparing the scores on the two interviews for each cue condition in Experiments II, III, and IV. Judgments on Interview B trials were more accurate than Interview A trials only under the body cue condition (Experiment III, $p < .02$; Experiment IV, $p < .001$). The results of Experiment IV on differences between judgments made on the basis of head or body cues provide related information. Head judgments were superior to body position judgments for Interview A ($p < .001$), while a trend in the reverse direction was seen for judgments made of Interview B.

There are a number of hypotheses which might explain why in one of the interviews head but not body cues could be related to verbal behavior. These explanations involve differences in sending behavior, that is, the extent to which people may differ in the information which is conveyed to an observer through their nonverbal behavior. The experimenter and the subject in Interview A may have been head rather than body senders; or, they may have been sending information in their bodies which was unrelated or contradictory to the verbal behavior; or, the body cues while related to the verbal behavior may have either anticipated it or lagged behind it. The choice between these hypotheses will have to await further research.

A communication paradigm involves consideration not only of sending behavior but also of receiving behavior, that is, possible differences across observers in their skill or sensitivity in understanding nonverbal behavior. While these experiments were not designed to study this question, one aspect of receiving behavior can be analyzed by determining whether accuracy in judging one interview was related to accuracy in judging the other. Spearman rank-order correlations were calculated between accuracy in judging Interview A and Interview B for the various cue conditions in the last three experiments. Of the six possible correlation coefficients, only two were significant and these were both in a negative direction. Thus, it would seem that accuracy in matching verbal and nonverbal behavior from one interview was independent of such accuracy in working with materials from another interview. The more general question of whether sensitivity as a receiver is specific to the behavior of a particular sender, or group of senders, must also await further research.

The overall results on the four experiments suggest that some information related to the verbal behavior is conveyed by spontaneous nonverbal behavior during interviews. In attempting to specify just what may be communicated through body position and facial expression, both specific and general classes of information can be outlined. The most obvious example of specific nonverbal communication would be a gesture such as a smile or fist shake which has a direct verbal equivalent or translation. Even less symbolic nonverbal acts such as swaying of the body or tapping of the foot may have a specific communicative value by emphasizing or focusing attention on a particular part of a verbal message. Although not studied in these experiments, nonverbal acts, such as movements towards or away from another or direction of eye gaze, may not necessarily be related to the verbal message, but instead may communicate specific information in the language of the relationship (Ekman, 1962). In addition to these specific meanings, nonverbal behavior may also communicate more general or gross information about the sender. Examples would be information about ac-

tivity level, anxiety, or the accumulation and discharge of tension.

In the present experiments, specific information must have been communicated by the nonverbal behavior. The judges were not simply responding to the gross differences between the behavior shown under stress and catharsis, since they had been able to discriminate between photograph pairs taken from the same emotional phase of the interview. This discrimination between two stress or two catharsis pictures was as accurate as the judgments made of pictures from different interview phases.

Although the majority of the results obtained were statistically significant, and the distribution of scores was skewed towards accuracy, the absolute level of correct choices was far from perfect. There were a number of factors which made this judgmental task extremely difficult. If the nonverbal behavior photographed had contradicted or was not directly related to the verbal behavior, accuracy would suffer. Similarly, if the nonverbal behavior anticipated or lagged behind the verbal behavior, the photograph could not have been matched with the short speech sample provided. Presenting the speech sample as a typescript without any vocal cues may have served to increase the ambiguity of the verbal behavior. Finally, the use of still photographs rather than motion pictures eliminated any cues from sequence or movement patterns which are customarily available to an observer.

In light of these limiting factors, it is particularly impressive that consistent evi-

dence was found for at least a partial relationship between verbal and nonverbal behavior. These experiments indicate that body position and facial expression spontaneously shown during an interview are not random activity or noise, but have specific communicative value related to the verbal behavior. Furthermore, this relationship is not obscure or available to only the privileged few, but can be detected by untrained observers.

REFERENCES

- DITTMANN, A. T. The relationship between body movements and moods in interviews. *J. consult. Psychol.*, 1962, 26, 480.
- EKMAN, P. Body language during interviews. Paper read at Inter-American Congress of Psychology, Mexico City, December 1961.
- EKMAN, P. The communication of interview stress through body language. Paper read at Western Psychological Association, San Francisco, April 1962.
- GIEDT, F. H. Comparison of visual, content, and auditory cues in interviewing. *J. consult. Psychol.*, 1955, 19, 407-416.
- MAHL, G., DANET, B., & NORTON, NEA. Reflection of major personality characteristics in gestures and body movements. *Amer. Psychologist*, 1959, 14, 357. (Abstract)
- SAINESBURY, P. Gestural movement during psychiatric interview. *Psychosom. Med.*, 1955, 17, 458-469.
- SCHLOSBERG, H. Three dimensions of emotion. *Psychol. Rev.*, 1954, 61, 81-88.
- SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
- VERNON, P. E. The matching method applied to investigations of personality. *Psychol. Bull.*, 1936, 33, 149-177.
- WOLFF, W. *The expression of personality*. New York: Harper, 1943.

(Received November 19, 1962)