Journal of Personality and Social Psychology
1982, Vol. 43, No. 2, 347-357

0022-3514/82/4302-0347$00.75

Face and Tone of Voice in the Communication of Deception

Miron Zuckerman, Mary D. Amidon, Shawn E. Bishop,
and Scott D. Pomerantz
University of Rochester

The contributions of face and tone of voice (filtered speech) to the communication of honest and deceptive messages were examined. In general, tone of voice was a better source of deception and leakage than the face. In addition, raters' judgments of the combined audiovisual channel were better predicted from their judgments of tone of voice when the message was deceptive and from their judgments of the face when the message was honest. The relative importance of face and tone of voice was also affected by the availability of verbal content—when content was available the face became less important. Thus, judges obtained more information from facial cues that were added to filtered speech (a comparison between filtered speech and face plus filtered speech) than from facial cues that were added to the full voice (a comparison between the voice and face plus voice). In addition, judgments of the audiovisual channel without content (face plus filtered speech) were better predicted from judgments of the face, whereas judgments of the audiovisual channel with content (face plus full voice) were better predicted from judgments of filtered speech. Finally, the relative importance of face and tone of voice was also determined by the affect that was communicated. Tone of voice was a better source of information about dominance and submission; the face revealed more information and was more highly correlated with the combined audiovisual channel for communications of liking and disliking.

The present study investigated the relative importance of face and tone of voice in the communication of honest and deceptive messages; these messages concerned the sender's feelings about four target persons: someone they liked, someone they disliked, someone they dominated, and someone to whom they submitted (cf. DePaulo & Rosenthal, 1979). Our choice of channels (face vs. tone of voice) and type of communication (truth/deception) was determined by two considerations. The first pertains to the "leakage" hierarchy of nonverbal channels (cf. Ekman & Friesen, 1969) and the second to the question of "video vs. audio primacy" (DePaulo, Rosenthal, Eisenstat, Rogers, & Finkelstein, 1978; Mehrabian & Ferris, 1967). These considerations are reviewed below. The choice of the target persons was based on the consideration that the face might be a better source of information about liking and disliking, whereas the tone of voice might be a better source of information about dominance and submission (Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979).

This research was supported by Naval Research Grant CNA SUB N00014-76-C 0001. The authors wish to thank Robert Rosenthal for his comments on an earlier version of this manuscript.

Requests for reprints should be sent to Miron Zuckerman, Department of Psychology, University of Rochester, Rochester, New York 14627.

The Leakage Hierarchy

Ekman and Friesen's (1969) concept of leakage hierarchy implies that some nonverbal channels are more controllable than others. Specifically, these investigators suggested that relative to the face, the body is less controllable and therefore a better source of leakage (nonverbal acts that give away a message the sender wishes to conceal) and deception cues (nonverbal acts indicating that deception is occurring without revealing the concealed message). Zuckerman, DePaulo, and Rosenthal (1981) conducted a quantitative summary of four studies of face–body differences in detectability of deception. In accordance with Ekman and Friesen's (1969) model and their original study (1974), it was found that the body is

more likely to reveal deception than is the face, $z = 4.48$, $p < .001$.

Recently, researchers began to extend Ekman and Friesen's leakage hierarchy to other nonverbal channels, particularly tone of voice (cf. Rosenthal, & DePaulo, 1979). Studies by Weitz (1972) and Bugental and her colleagues (Bugental, Caporael, & Shennum, 1980; Bugental, Henker, & Whalen, 1976; Bugental & Love, 1975) indicated that tone of voice leaks information that is not revealed by the more controllable verbal content. A related study (Zuckerman, Larrance, Spiegel, & Klorman, 1981) indicated that like the body, tone of voice is less controllable than facial expressions. Specifically, Zuckerman et al. found that senders were better able to modify (suppress and exaggerate) facial expressions than tone of voice. Senders were also more aware of their facial-sending skill than of their vocal-sending skill. Finally, sending accuracy under suppression (i.e., the information that was leaked in spite of instructions to suppress expressions) was more highly correlated with sending accuracy under exaggeration for vocal than for facial expressions. Interpretation of this finding was that since the voice is a leaky channel, what it leaks under suppression is a good predictor of what it sends under exaggeration; since the face is controllable, what it leaks under suppression is not a good predictor of what it sends under exaggeration.

To the extent that tone of voice is less controllable than the face, tone of voice should be a better source of deception and leakage cues. This prediction was examined in the present study.

Video Versus Audio Primacy

The question of video versus audio primacy concerns the relative influence of the visual and auditory channels on perceiver's judgments. Several studies (Bugental, Kaswan, & Love, 1970; DePaulo, Rosenthal, Eisenstat, Rogers, & Finkelstein, 1978; Mehrabian & Ferris, 1967) suggested that in the resolution of conflicting video and audio messages, the face receives greater weight than does the tone of voice. For example, DePaulo et al. (1978) showed that a positive facial expression combined with a negative tone of voice was rated more positively than a negative facial expression combined with a positive tone of voice. Stated differently, judges' impressions of the discrepant messages were more influenced by the face than by the tone of voice. However, more recent studies (Ekman, Friesen, O'Sullivan, & Scherer, 1980; Krauss, Apple, Morency, Wenzel, & Winton, 1981) showed that in general, correlations of judgments of tone of voice or of voice (words plus tone) with judgments of the full audiovisual channel were not lower than the corresponding correlations of the face with the full channel. Both Ekman et al. and Krauss et al. concluded that the relative influence of different channels may depend on several factors including the affect that is being judged, the situation in which the message was elicited, and various apects of the communication (e.g., degree of consistency among the competing channels).

One potentially important aspect of the communication concerns the availability of verbal content. In two of the studies supporting video primacy (DePaulo et al., 1978; Mehrabian & Ferris, 1967), the observers responded to presentations of face plus *only* tone of voice; content was eliminated by the use of content filtering (DePaulo et al., 1978), randomized splicing (DePaulo et al., 1978), or a single word standard speech (Mehrabian & Ferris, 1967). The Bugental et al. (1976) study is a special case. Although full speech was used, it was made of either consistent or conflicting content and tone. Face received greater weight when compared to either tone of voice or content alone. However, face received smaller weight when compared to consistent speech (content plus tone of voice). Specifically, a positive facial expression combined with a negative speech (negative content plus negative tone of voice) was rated *less* positively than a negative facial expression combined with a positive speech. It appears that the greater weight of the face relative to the tone of voice was obtained only when the latter could not be assisted by content. A similar inference can be drawn from the two studies that did not support video primacy (Ekman et al., 1980; Krauss et al., 1981). In both studies,

observers responded to presentations of face plus full voice. Perhaps observers pay more attention to the tone of a verbally intelligible message than to the tone of either verbally meaningless sound or inconsistent content. Stated differently, the face may weigh more heavily relative to tone of voice when content of speech is absent or conflicting.

As a preliminary test of the above notion, we examined what observers learned from facial expressions that were added to filtered speech (a comparison between filtered speech and face plus filtered speech) and from facial expressions that were added to full voice (a comparison between full voice and face plus full voice). Addition of facial cues to filtered speech was expected to produce a greater increase in perceivers' accuracy than the addition of facial cues to the full voice.

Correlations of Face and Tone of Voice With the Audiovisual Channel

In the previous sections we proposed to examine the communication accuracy of various combinations of facial and vocal cues as a means of testing ideas regarding the nonverbal leakage hierarchy and video versus audio primacy. Communication accuracy reflects, of course, the amount of information revealed by the senders and acquired by the receivers. As noted by Krauss et al. (1981) and Ekman et al. (1980), it is also possible to examine the correlations between judgments of each channel and judgments of the combined audiovisual record. Such correlations estimate each channel's contribution to the perception of the combined channel presentation. In the present study, these correlations served as additional and somewhat different tests of our ideas regarding the nonverbal leakage hierarchy and the video versus audio primacy.

To the extent that tone of voice is leakier than the face, perceivers expecting deception may pay more attention to the auditory component of the message. In a recent study, Zuckerman, Spiegel, DePaulo, and Rosenthal (in press) administered to subjects a video primacy test—an instrument composed of multiple channel visual (face or body) plus auditory cues (content filtered or random spliced) that are either consistent or discrepant (cf. DePaulo et al., 1978). It was found that subjects who were told that the stimulus person in the test lied were less influenced by the face relative to the voice than subjects who were told that the person never lied. In the previously mentioned Ekman et al. (1980) study, judgments of the full audiovisual presentation of deceptive messages were more highly correlated with judgments of the voice than with judgments of either the face or body; when honest messages were examined, there were no differences among the correlations of the three channels with the criterion. The judges in this study were not told that some of the messages involved deception. Perhaps changes in pitch or other voice qualities associated with deception drew attention to the vocal channel when senders were lying (Ekman et al., 1980, p. 27).

It thus appears that either expectations of deception or vocal qualities associated with deception may draw the observer's attention away from the face and onto the voice. Accordingly, we expected that correlations between tone of voice and the combined audiovisual channel would be higher for deception than for honest messages; correlations between the face and the combined channel were expected to show the opposite pattern.

Predictions about correlations between the separate channels and the combined channel can also be derived from our previous analysis of the video primacy issue. Earlier we speculated that judges pay more attention to tone of voice associated with meaningful verbal content than to tone of voice without content. Accordingly, it can be expected that judgment of filtered speech will be more highly correlated with judgment of face plus voice than with judgment of face plus tone of voice; correlations of the face with these two audiovisual presentations might show the opposite pattern.

In summary, the present study examined two notions regarding tone of voice and facial expressions. The first was that relative to the face, tone of voice is less controllable and hence (a) provides more deception and leakage cues and (b) attracts more attention when the message is deceptive. The second was that the face attracts more attention when the message has no content and hence

(a) provides more information when added to tone of voice than to voice and (b) exerts more influence on judgments of face plus tone of voice than on judgments of face plus voice.

Method

Subjects

Subject senders were 30 male and 30 female undergraduates who participated in partial fulfillment of a research requirement for an introductory psychology course.

Sending Session

The subject senders sat in a comfortable armchair and faced the experimenter, who sat 7 feet (2.13 m) away, and a videotape camera located at the right side of the experimenter. Equipment operated by an assistant in an adjacent room permitted videotaping, audiotaping, and on-line monitoring of the subject's reactions.

The procedure was similar to the one used by DePaulo and Rosenthal (1979). Subject senders described four target persons of their choice: someone they liked (a "like" target), someone they disliked (a "dislike" target), someone they dominated, and someone to whom they submitted ("dominate" and "submit" targets, respectively). Each target person was described in three modes—truth, concealment, and deception—resulting in 12 descriptions for each sender (4 Targets × 3 Modes). In the truthful mode, the senders were asked to convince the experimenter about their true feelings toward the target; in the concealment mode, the senders were asked to describe the target in a way that would not reveal how they felt about this person; and in the deceptive mode, they were asked to convince the experimenter that their feelings about the target were opposite to how they actually felt, that is, that they disliked the like target, liked the dislike target, submitted to the dominate target, and dominated the submit target. It was thought that the inclusion of four affects (liking–disliking and dominance–submission) and three sending modes would increase the external validity of the study.

Subjects gave each description in one of 12 sequences that formed a Latin square, with the restriction that each block of four descriptions in each sequence included all four target persons. The experiment was presented as a study of communication skills, and subjects were urged to perform as well as they could. They were given 1 min for each description and asked to talk during the entire time period. The middle 35 sec of each description was videotaped and audiotaped.

Sending Scores

The recorded stimulus materials were presented to nine paid undergraduates judges (five females and four males) in one of five channels:

1. *Face.* Judges saw the sender's facial expressions without the soundtrack.
2. *Tone of voice (filtered speech).* Judges heard the sound track after it had been electronically filtered (cf. Rogers, Scherer, & Rosenthal, 1971; Starkweather, 1956), a process that renders the speech content unintelligible but leaves most expressive vocal qualities intact; the face was not shown.
3. *Voice.* Judges heard the unaltered sound track (words plus tone of voice) but did not see the senders.
4. *Face plus voice.* Judges observed the facial expressions plus the unaltered sound track associated with each message.
5. *Face plus tone of voice.* Judges observed the facial expressions plus the filtered speech associated with each message.

Each judge rated all the stimulus materials in all channels, participating in about 30 rating sessions that spanned a period of a month and a half. Each rating session lasted 1½ hours and was devoted to the materials of 9 to 11 senders in a single channel. Assignment of channels and senders to rating sessions was random, with the restrictions that each group of five consecutive sessions covered all five channels and each group of six rating sessions covered all 60 senders.

Judges were informed that in each description the senders either told the truth; concealed the truth; or lied about someone they either liked, disliked, dominated, or submitted to. The judges rated each description on two 9-point scales, liking (1 = sender dislikes this person, 9 = sender likes this person) and dominance (1 = sender submits to this person, 9 = sender dominates this person); in addition, they noted whether the description in question involved deception, concealment, or truth. It was stressed to the judges at the beginning of almost every session that their ratings should reflect the senders' actual affects rather than the affects the senders tried to project. For example, a sender pretending to like someone that he or she actually disliked was to be given a low score on the liking scale.

Senders were used as units of analysis. Both liking and dominance ratings were averaged across judges, yielding two affective scores for each of the 12 descriptions. In addition, the proportion of judges who correctly identified a description as either deception, concealment, or truth was also computed, yielding deception detection accuracy scores that could range from 0 to 1.00, with .33 as the level of chance accuracy. In summary, each of the 12 descriptions were assigned three scores—liking, dominance, and detection accuracy.

Results

Face Versus Filtered Speech: Deception and Leakage Information

To determine which channel was a better source of deception cues, we examined the deception detection accuracy scores in an analysis of variance with sex of sender as a between-subjects factor[1] and channel (face

[1] Sex of senders was included in all analyses of variance. However, gender did not affect any of the results presented in this article and so is not discussed.

Table 1
Deception Detection Accuracy of Filtered Speech and Face by Affect

Channel	Affect				
	Liking	Disliking	Dominance	Submission	*M*
Filtered speech	.43	.40	.42	.41	.41[a]
Face	.41	.39	.36	.37	.38[b]
Difference	.02	.01	.06	.04	.03

[a] Significantly greater than chance, $F(1, 58) = 10.24$, $p < .001$, $r = .39$.
[b] Slightly greater than chance, $F(1, 58) = 3.06$, $p < .10$, $r = .24$.

vs. filtered speech), affect (liking/disliking/dominance/submission), and mode (truth/concealment/deception) as within-subjects factors. Table 1 presents mean detection deception accuracy scores for each channel and affect.[2] It can be seen that accuracy was higher for filtered speech than for the face, $F(1, 58) = 13.75$, $p < .001$, $r = .44$.[3] It is interesting that the difference between the two channels was greater for dominance and submission than for liking and disliking; the interaction of channel with the two pairs of affects (average of liking and disliking vs. average of dominance and submission) approached significance, $F(1, 58) = 3.45$, $p < .07$, $r = .24$.

To determine which channel was a better source of leakage, we examined the liking and dominance ratings in two separate analyses of variance, each with target person (like/dislike or dominate/submit) as a within-subjects factor; in both analyses, sex was a between-subjects factor and sending mode (truth/concealment/deception) a within-subjects factor. Table 2 presents the relevant mean scores. It can be seen that, in general, liking scores were higher for the like than for the dislike target person, $F(1, 58) = 22.25$, $p < .001$, $r = .53$. Although the difference between the like and dislike targets, which can be conceptualized as the amount of information revealed by the channel, was slightly higher for filtered speech, the interaction between target and channel was not significant ($F < 1$).

Dominance scores were higher for the dominate target person than for the submit target, $F(1, 58) = 39.20$, $p < .001$, $r = .64$, a difference that was greater for filtered speech than for face, Channel × Target, $F(1, 58) = 5.96$, $p < .02$, $r = .31$. Note that filtered speech was more informative than the face in deception as well as in truth. Why was this difference not limited to the deception mode? It seems that the judges, warned to rate the senders' true feelings, could not take any expression at "face" value. Since they were less able to distinguish truth from deception in the face than in the tone of voice, they were more likely to suspect deception and therefore to discount truthful facial cues. Thus, they obtained less information from the face relative to tone of voice in all communication modes rather than in only the deception mode.

[2] Differences in detection accuracy among the three modes are not presented because they may reflect biases that are irrelevant to the purpose of the present study. Thus, a particularly high level of detection accuracy in the truth mode may simple indicate (a) that senders tended to appear honest and/or (b) that judges tended to interpret messages as honest regardless of the actual content of the messages. An interaction between mode and channel may simply indicate that channels vary in the extent to which they give rise to either of the above biases. These biases are avoided when the overall detection accuracy (averaged across modes) is examined, because the bias that increases the accuracy in one mode will decrease it in another. (For a more elaborate discusssion of these issues, see Zuckerman, DePaulo, and Rosenthal, 1981, and Zuckerman, DeFrank, Hall, Larrance, and Rosenthal, 1979.)

[3] An estimate of effect size, r is computed as

$$\sqrt{\frac{F}{df_e + F}}$$

when there is only one degree of freedom for the numerator of F. The relationship of r with Cohen's (1977) d, which is another estimate of effect size, is given by the equation

$$r = \frac{d}{\sqrt{d^2 + 4}}.$$

Overall, the results support the hypothesis that tone of voice is a leakier channel than the face, but the difference was smaller for communication of liking and disliking. This latter finding is consistent with the notion that tone of voice provides more information about dominance and submission, whereas the face provides more information about liking and disliking.

Effects of Availability of Verbal Content on Facial Information

To examine the amount of information that judges learned from the face in the presence and absence of verbal content, we conducted separate analyses of variance on deception detection accuracy, liking ratings, and dominance ratings. In all three analyses, face (present/absent) was crossed with verbal content (present/absent), resulting in the factorial design presented in Table 3. It can be seen that the two main effects in this design indicate the amount of information that judges obtain from the addition of facial cues and from the addition of content to the message. More important, the Face × Content interaction indicates whether the increase in information due to the addition of facial cues is greater when content is absent (a comparison between filtered speech and face plus filtered speech) than when content is present (a comparison between voice and face plus voice).

Table 2
Information Accuracy of Filtered Speech and Face by Affect and Mode of Communication

Channel (affect)	Truth	Concealment	Deception	*M*
Liking ratings				
Filtered speech				
Like	5.48	5.53	5.16	5.39
Dislike	4.80	5.41	5.02	5.08
Difference	.68	.12	.14	.31
Face				
Like	5.66	5.26	5.23	5.38
Dislike	5.10	5.10	5.20	5.13
Difference	.56	.16	.03	.25
Dominance ratings				
Filtered speech				
Dominate	5.18	5.10	5.10	5.13
Submit	4.47	4.77	4.57	4.60
Difference	.71	.33	.53	.53
Face				
Dominate	4.96	5.10	5.03	5.03
Submit	4.83	4.86	4.74	4.81
Difference	.13	.24	.29	.22

Table 3
Classification of Four Sending Modes as a Function of the Presence/Absence of Face and Content

	Content	
Face	Present	Absent
Present	Face plus voice	Face plus filtered speech
Absent	Voice	Filtered speech

In the analysis of deception detection accuracy, affect and mode were added to face and content as the within-subjects factors; sex was a between-subjects factor. Table 4 presents the mean accuracy scores for each communication channel and affect. It can be seen that adding facial cues increased detection accuracy, $F(1, 58) = 18.81, p < .001$, $r = .49$, particularly for liking and disliking: interaction of Face × Pairs of Affects (liking and disliking versus dominance and submission), $F(1, 58) = 4.07, p < .05, r = .26$. Similarly, adding content also increased detection accuracy, $F(1, 58) = 69.70, p < .001$, $r = .74$. Furthermore, the increase in accuracy due to the addition of facial cues was somewhat greater when content of speech was absent, Face × Content, $F(1, 58) = 2.88, p < .10, r = .22$.

In the analyses of the liking and dominance ratings, face, voice, sending mode, and target (either like/dislike or dominate/submit) were the within-subjects factors and sex the between-subjects factor. To facilitate presentation, in Table 5 we show the difference in ratings between targets, that is., liking ratings of the like target minus liking ratings of the dislike target, and dominance ratings of the dominate target minus dominance ratings of the submit target; the

Table 4
Deception Detection Accuracy of Four Channels by Affect

Face	Liking		Disliking		Dominance		Submission	
	Content present	Content absent	Content present	Content absent	Content present	Content absent	Content present	Content absent
Present	.58	.50	.51	.46	.48	.44	.52	.46
Absent	.54	.43	.48	.40	.48	.42	.49	.41
Difference	.04	.07	.03	.06	.00	.02	.03	.05

Note. Accuracy was significantly above chance for all channel combinations ($ps < .001$).

greater the difference, the greater the amount of information revealed by the channel.

It can be seen that all liking difference scores were positive, indicating that the like target received greater liking ratings than the dislike target, $F(1, 58) = 43.81, p < .001, r = .66$. This difference was greater when facial cues were added to the message, Face × Target, $F(1, 58) = 4.44, p < .05, r = .27$, and when content was added to the message, Content × Target, $F(1, 58) = 11.62, p < .002, r = .41$. More important, the increase in revealed information due to the addition of facial cues was greater when content was absent than when it was present, Face × Voice × Target, $F(1, 58) = 8.21, p < .006, r = .35$; this three-way interaction was not affected by mode of sending ($F < 1$).

Similarly, it can be seen that all dominance difference scores were positive, indicating that the dominate target received greater dominance ratings than the submit target, $F(1, 58) = 96.33, p < .001, r = 79$. This difference was only slightly greater when facial cues were added to the message, Face × Target, $F(1, 58) = 2.22, p < .15, r = .19$, and significantly greater when content was added to the message, Content × Target, $F(1, 58) = 35.28, p < .001, r = .61$. More important, the increase in revealed information due to the addition of facial cues was greater when content was absent than when it was present, Face × Content × Target, $F(1, 58) = 5.28, p < .03, r = .29$; this three-way interaction was somewhat greater in the truth mode relative to the two other modes, Face × Content × Target × Mode, $F(2, 116) = 2.89, p < .07, \eta = .30$.

In summary, the three analyses of variance indicated that judges obtained more information from the face when the content of the message was not available.

Correlations of Face and Filtered Speech With the Audiovisual Channel

Correlations of judgments of face and filtered speech with judgments of combined audiovisual channel were examined at each level of the following factors (names of factors are followed by names of levels in each

Table 5
Information Accuracy of Four Channels by Affect

Face	Liking–disliking[a]		Dominance–submission[b]	
	Content present	Content absent	Content present	Content absent
Present	.85	.74	1.29	.82
Absent	.84	.31	1.39	.52
Difference	.01	.43	−.10	.30

[a] Liking ratings of "like" target minus liking rating of "dislike" target.
[b] Dominance ratings of "dominate" target minus dominance ratings of "submit" target.

Table 6
Mean Correlations of Face and Filtered Speech With the Combined Channels for Truth, Concealment, and Deception

Mode	Channel		
	Face	Filtered speech	Difference
Truth	.23	.15	.08
Concealment	.21	.23	−.02
Deception	.19	.26	−.07

factor): (a) combined channel: face and voice, face and filtered speech; (b) affect: liking, disliking, dominance, submission; (c) mode: truth, concealment, deception; (d) type of scores: deception detection accuracy scores, liking and dominance "relevant," liking and dominance "irrelevant."[4]

Since the factors were fully crossed, there was a total of 144 correlations of face and filtered speech with the combined channels, one correlation in each cell of the following factorial design: 2 Channels (face vs. filtered speech) × 2 Combined Channels × 4 Affects × 3 Modes × 3 Types of Scores. *Z* transformations of these correlations were examined in an analysis of variance in which all the above factors served as within-subjects factors with $n = 1$ (only one correlation in each cell). Mean correlations of two two-way interactions, Channel × Mode and Channel × Combined Channel, were of interest to the experimental hypotheses; they are presented in Tables 6 and 7. Each mean correlation is the average correlation, based on *z* transformations, across all factors not presented in the interaction.

Table 6 indicates that mean correlations of tone of voice with the combined channels increased from truth to concealment and from concealment to deception; mean correlations of face with the combined channels showed the opposite pattern. To examine the significance and, more important, the effect size of the Channel × Linear Mode interaction, a residual error term was constructed by aggregating the sum of squares of all the three-way and higher order interactions and dividing the results by the sum of the *df*s associated with these interactions (Green & Tukey, 1960). The sum of squares of the Channel × Linear Mode was divided by this pooled error, resulting in $F(1, 103) = 11.78$, $p < .001$, $r = .32$. It should be noted that this procedure is biased in the direction of yielding too small *F*s because it uses an estimate of the error term that might include nonerror components and consequently may be larger than the true error.

Table 7
Mean Correlations of Face and Filtered Speech With Two Combined Channels

Combined channel	Channel	
	Face	Filtered speech
Face plus voice	.18	.25
Face plus filtered speech	.24	.18

Table 7 shows the mean correlations of face and filtered speech with two types of combined channel. It can be seen that filtered speech was more highly correlated with face plus voice than with face plus filtered speech; correlations of the face with the two combined channels showed the opposite pattern. The procedure of constructing an error term by aggregation yielded a significant Channel (face vs. filtered speech) × Combined Channel (face plus voice vs. face plus filtered speech) interaction, $F(1, 103) = 12.65$, $p < .001$, $r = .33$.

Since the previous results indicated that the face reveals more information about liking and disliking than about dominance and submission, we present the mean correlations of face and filtered speech with the combined channels for each affect (see Table 8). It appears that judgments of the face were more highly correlated with judgments of combined channels for liking and disliking

[4] Liking and dominance "relevant" refer to ratings that were used in all previous analyses, that is, the liking ratings of the like and dislike target persons and the dominance ratings of the dominant and submit targets. The liking and dominance "irrelevant" refer to the liking ratings of the dominate and submit targets and the dominance ratings of the like and dislike targets (recall that the judges rated *all* targets on both liking and dominance); these latter ratings are irrelevant in that they reflect target's attributes that the senders did not try to reveal, conceal, or lie about. Although these ratings were irrelevant to analyses of the amount of information revealed by the senders, they could be included in analyses of the relationship between judgment of a single channel and judgments of the combined channel.

than for dominance and submission; correlations of the filtered speech with the combined channels did not vary as a function of affect. The procedure of constructing an error term by aggregation produced a significant interaction between channel (face vs. filtered speech) and the two pairs of affects (liking and disliking vs. dominance and submission), $F(1, 103) = 6.93$, $p < .01$, $r = .25$.

In general, the results reported in this section are consistent with the previous findings regarding communication accuracy: The leakier voice exerted more influence on perception of combined channel deceptive communications, the face exerted more influence on perception of combined channel without content, and the face also exerted more influence on perception of combined channel expressions of liking and disliking.

Table 8
Mean Correlations of Face and Filtered Speech With the Combined Channels for Four Affects

	Channel	
Affect	Face	Filtered speech
Liking	.26	.18
Disliking	.27	.25
Dominance	.20	.22
Submission	.12	.21

Discussion

In recent years there has been a marked increase in the volume of research on nonverbal communication, producing changes in the type of questions asked, the complexity of experimental procedures, and the type of responses that are expected. More specifically, earlier claims of the primacy of nonverbal cues (relative to verbal cues) and of facial expressions (relative to auditory cues) were changed into the hypothesis that the relative importance of a specific channel would depend on several variables, including "the sort of judgment the subject is asked to make, the particular setting, and even the configuration of cues among channels" (Krauss et al., 1981, p. 319). Furthermore, earlier procedures designed to examine the communication accuracy of separate channels now coexist with procedures designed to examine each channel's contribution to judgments of the combined audiovisual presentation, regardless of the channels' communication accuracy. Naturally, the proliferation of experimental procedures has also led to discussions of their theoretical underpinnings and of the consistency between results that the different methods produce.

The present study examined the communication accuracy of face and tone of voice as well as their correlations with the combined audiovisual record. The hypotheses tested concerned two issues: the leakage hierarchy model and the effects of content availability on video primacy. Of the four affects that were examined, two (liking and disliking) were considered more pertinent to the face and two (dominance and submission) were considered more pertinent to the voice. The results indicated that all three factors—leakiness of channels, availability of content, and affective content of the message—accounted for some of the variance in the communication process. Each of these factors will be reviewed separately.

Generally speaking, the results showed that tone of voice is a better source of deception and leakage cues than is the face. The difference in actual leakiness between the two channels raises the question of difference in their perceived leakiness; that is, is the actually leakier channel also treated as such by receivers of the communication? The investigation of whether receivers treat tone of voice as a leaky channel must be based on a theory that incorporates perceived leakiness as one of its variables. The theory tested here was that perceivers expecting deception would be more influenced by leaky cues (tone of voice) than by controllable cues (facial expressions) that accompany deceptive messages. The results supported these predictions—tone of voice was more likely to influence judgments of deceptive messages, whereas face was more likely to influence judgments of truthful messages (see Table 6).

Note that perceived leakiness of face and tone of voice were established here according to their relative contributions to the perception of the combined channel and not according to what judges said about the leakiness of these channels. Whether judges' ver-

bal reports about leakiness of face and tone of voice correspond to either actual or perceived leakiness remains to be examined in future research (cf. Zuckerman, Koestner, & Driver, 1981).

Video primacy was examined as a function of the availability of verbal content. It was found that the face is a better source of information in the absence of verbal content. First, judges obtained more information from facial cues added to filtered speech than from facial cues added to full voice. Second, judgments of face plus filtered speech were more highly correlated with judgments of face than with judgments of tone of voice; correlations of face plus voice with the two channels showed the opposite pattern. These findings may have both methodological and theoretical implications.

From a methodological point of view, the results suggest that previous demonstrations of video primacy—the findings that observers responding to face plus tone of voice were more influenced by the face than by the tone of voice (Bugental et al., 1970; DePaulo et al., 1978; Mehrabian & Ferris, 1967)—may not generalize to situations in which observers respond to naturally occurring messages with content (cf. Krauss et al., 1981). In addition, the results demonstrate the active role of observers in the communication process. Thus, when there was no verbal content to draw their attention, judges obtained more information from the face. Perhaps the importance of various channels can be determined by specific instructions or training that would lead judges to focus on a particular channel (cf. DePaulo, Lassiter, & Stone, in press).

Turning to the question of message content, both communication accuracy and the relationship of face and tone of voice with the combined audiovideo channel were influenced by the type of affect that was communicated. Thus, the greater leakiness of the voice was more emphasized for dominance and submission than for liking and disliking (see Tables 1 and 2); the increase in accuracy of the communication due to the addition of facial cues were more emphasized for liking and disliking than for dominance and submission (see Tables 4 and 5); and the judgments of the face were more highly correlated with judgments of the audiovisual channel for liking and disliking than for dominance and submission (see Table 8). In general, these results are consistent with previous suggestions that the face is a better source of information about the liking dimension, whereas the voice communicates better feelings about dominance and submission (DePaulo et al., 1978; Ekman et al., 1980; Rosenthal et al., 1979; Zuckerman et al., 1979).

It seems appropriate to end with a short note on the ecological validity of the study. The types of communication that were examined are closely related to real-life events. People often tell the truth, conceal information, or lie about their feelings toward both friends and enemies. The context of the communication was, of course, highly artificial but no more than that of a standard experiment on psychophysiological detection of lying. There is a theoretical link between physiological and nonverbal detection of deception (Hemsley, 1977); in fact, the same processes may account for both the physiological and the nonverbal cues accompanying deception (cf. Zuckerman et al., 1981). It is, therefore, interesting to note that laboratory research on physiological lie detection generalizes with some degree of success to field setting. Although it is essentially an empirical question, there may be no need to despair over the artificiality of research on nonverbal lie detection and the limits of its applicability.

References

Bugental, D. B., Henker, B., & Whalen, C. K. Attributional antecedents of verbal and vocal assertiveness. *Journal of Personality and Social Psychology*, 1976, *34*, 405–411.

Bugental, D. B., & Love, L. Nonassertive expression of parental approval and disapproval and its relationship to child disturbance. *Child Development*, 1975, *46*, 747–752.

Bugental, D. E., Caporael, L., & Shennum, W. A. Experimentally-produced child uncontrollability: Effects on the potency of adult communication patterns. *Child Development*, 1980, *51*, 520–528.

Bugental, D. E., Kaswan, J. W., & Love, L. R. Perception of contradictory meanings conveyed by verbal and nonverbal channels. *Journal of Personality and Social Psychology*, 1970, *16*, 647–655.

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

DePaulo, B. M., Lassiter, G. D., & Stone, J. I. Attentional determinants of success at detecting deception. *Personality and Social Psychology Bulletin*, in press.

DePaulo, B. M., & Rosenthal, R. Telling lies. *Journal of Personality and Social Psychology*, 1979, *37*, 1713–1722.

DePaulo, B. M., Rosenthal, R., Eisenstat, R. A., Rogers, P. L., & Finkelstein, S. Decoding discrepant nonverbal cues. *Journal of Personality and Social Psychology*, 1978, *36*, 313–323.

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

Ekman, P., & Friesen, W. V. Detecting deception from the body or face. *Journal of Personality and Social Psychology*, 1974, *29*, 288–298.

Ekman, P., Friesen, W. V., O'Sullivan, M., & Scherer, K. Relative importance of face, body, and speech in judgments of personality and affect. *Journal of Personality and Social Psychology*, 1980, *38*, 270–277.

Green, B. F., & Tukey, J. W. Complex analyses of variance: General problems. *Psychometrica*, 1960, *25*, 127–152.

Hemsley, G. D. *Experimental studies in the behavioral indicants of deception.* Unpublished doctoral dissertation, University of Toronto, 1977.

Krauss, R. M., Apple, W., Morency, N., Wenzel, C., & Winton, W. Verbal, vocal, and visible factors in judgments of another's affect. *Journal of Personality and Social Psychology*, 1981, *40*, 312–319.

Mehrabian, A., & Ferris, S. R. Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology*, 1967, *31*, 248–252.

Rogers, P. L., Scherer, K. R., & Rosenthal, R. Content filtering human speech. A simple electronic system. *Behavior Research Methods and Instrumentation*, 1971, *3*, 16–18.

Rosenthal, R., & DePaulo, B. M. Sex differences in eavesdropping on nonverbal cues. *Journal of Personality and Social Psychology*, 1979, *37*, 273–285.

Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, P. L., & Archer, D. *Sensitivity to nonverbal communication: The PONS test.* Baltimore, Md.: Johns Hopkins University Press, 1979.

Starkweather, J. Content-free speech as a source of information about the speaker. *Journal of Abnormal and Social Psychology*, 1956, *52*, 394–402.

Weitz, S. Attitude, voice, and behavior. A repressed affect model of interracial interaction. *Journal of Personality and Social Psychology*, 1972, *24*, 12–21.

Zuckerman, M., DeFrank, R. S., Hall, J. A., Larrance, D. T., & Rosenthal, R. Facial and vocal cues of deception and honesty. *Journal of Experimental Social Psychology*, 1979, *15*, 378–396.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14). New York: Academic Press, 1981

Zuckerman, M., Koestner, R., & Driver, R. Beliefs about cues associated with deception. *Journal of Nonverbal Behavior*, 1981, *6*, 105–114.

Zuckerman, M., Larrance, D. T., Spiegel, N. H., & Klorman, R. Controlling nonverbal cues: Facial expressions and tone of voice. *Journal of Experimental Social Psychology*, 1981, *17*, 506–524.

Zuckerman, M., Spiegel, N. H., DePaulo. B. M., & Rosenthal, R. Nonverbal strategies for decoding deception. *Journal of Nonverbal Behavior*, in press.

Received July 6, 1981 ■