

## RELATIONSHIPS BETWEEN SPEECH DELIVERY AND SPEECH EFFECTIVENESS

D. F. GUNDERSEN and ROBERT HOPPER

MANY teachers believe that speech delivery can enhance speech effectiveness, but this proposition has not been adequately tested. On the basis of existing research it is not clear what behaviors constitute "good" speech delivery, nor is it clear in what ways or whether such behaviors relate to speech effectiveness. In fact, the nature of speech effectiveness is still somewhat in question.

Historically, the discipline has shown an interest in speech delivery and its importance to speech effectiveness. Brooks summarized the tradition: "Since the time of Aristotle, speech scholars and social scientists have attempted to discover the specific characteristics which enhance the speaker's communication effectiveness."<sup>1</sup> One of the five canons of rhetoric, *eleutio*, is delivery. Cicero reported that "when Demosthenes was asked what is the first thing in speaking, he assigned the first role to delivery, and also the second, and also the third." Cicero agreed: "Delivery, I assert, is the dominant factor in oratory."<sup>2</sup> Thompson, in reviewing over thirty years of research, suggested that delivery affects comprehension and persuasiveness significantly.<sup>3</sup> He indicated

a wide variation in the nature of the delivery variables manipulated in research and in the definitions of effectiveness in delivery studies.<sup>4</sup>

Petrie, who surveyed research concerned with informative speaking, found only equivocal relationships between delivery behaviors and various measures of effectiveness.<sup>5</sup> He concluded that variables such as poor vocal quality, pitch patterns, nonfluencies *seem* not to have any relationship to audience recall and comprehension; but gestures, pauses, raising of the voice, visual aids, and eye contact did seem to be related to such measures of effectiveness.<sup>6</sup>

Barker, Kibler, and Geter<sup>7</sup> indicated that more study was needed to determine the relationships between specific speech delivery behaviors and measures of effectiveness, such as attitude change and content recall. The research to be reported here investigated some of these relationships. The problem was divided into three parts: 1. What are those delivery behaviors which are generally agreed to be necessary to speech effectiveness, and can a subset of these be operationalized as independent variables? 2. What constitutes "speech effectiveness" and how can measures of

*Mr. Gundersen is Assistant Professor of Speech at the University of Alabama and Mr. Hopper is Associate Professor of Speech Communication at the University of Texas, Austin.*

<sup>1</sup> William D. Brooks, *Speech Communication* (Dubuque, Ia.: William C. Brown, 1971), pp. 161-62.

<sup>2</sup> Cicero, *De Oratore*, trans. H. Rackham (Cambridge, Mass.: Harvard Univ. Press, 1960), II, 169.

<sup>3</sup> Wayne N. Thompson, *Quantitative Re-*

*search in Public Address and Communication* (New York: Random House, 1967), p. 82.

<sup>4</sup> Thompson, p. 84.

<sup>5</sup> Charles R. Petrie, Jr., "Informative Speaking: A Summary and Bibliography of Related Research," *SM*, 30 (1963), 85.

<sup>6</sup> Petrie, p. 80.

<sup>7</sup> Larry L. Barker, Robert J. Kibler, and Rudolph W. Geter, "Two Investigations of the Relationship among Selected Ratings of Speech Effectiveness and Comprehension," *SM*, 35 (1968), 405-06.

effectiveness be operationalized as dependent variables? 3. What relationships can be demonstrated to exist between effective delivery behaviors and speech effectiveness?

STUDY I

*Delivery Behaviors*

In order to answer question one, the researchers examined eleven representative speech texts and four empirical studies<sup>8</sup> to determine which delivery be-

haviors were considered to be related to speech effectiveness. Six such delivery behaviors were listed by more than 70 percent of the authors: vocal volume, rate of speech, voice quality, posture, gesture, and body movement. The first three can be characterized as "verbal" behaviors; the last three as "nonverbal."

Table I lists the operational definitions used for the six delivery behaviors in this research. Each variable was oper-

*Guide to Good Speech* 3rd. ed. (Englewood Cliffs, N.J.: Prentice Hall, 1965); James C. McCroskey, *An Introduction to Rhetorical Communication* (Englewood Cliffs, N.J.: Prentice-Hall, 1968); Alan H. Monroe and Douglas Ehninger, *Principles and Types of Speech* (Glenview, Ill.: Scott, Foresman, 1967); Thomas Scheidel, *Speech Communication and Human Interaction* (Glenview, Ill.: Scott, Foresman, 1972); Erwin P. Bettinghaus, "The Operation of Congruity in an Oral Communication Situation," *SM*, 28 (1961), 131-42; W. K. Clark, "A Survey of Certain Audience Attitudes toward Commonly Taught Standards of Public Speaking," *SM*, 18 (1951), 62-69; Petrie, "Informative Speaking"; John W. Wills, "An Empirical Study of the Behavioral Characteristics of Sincere and Insincere Speakers," Diss. Univ. of Southern California, 1961.

TABLE I  
OPERATIONAL DEFINITIONS OF INDEPENDENT VARIABLES\*

Variable	Effective	Ineffective
Volume	Normal for Sony 3600. Automatic gain control.	Audio track attenuated 20 db at source.
Rate	150 wpm (+ or -5).	125 wpm (+ or -5)
Voice Quality	Pleasant, does not call attention to itself. Normal to speaker	Nasal, with vocal mechanisms tensed to produce harshness and upward pitch change in the voice.
Posture	Feet 12 inches apart, one slightly ahead of the other. Head erect, chin up, chest out, shoulders relaxed. Standing on balls of feet.	Feet 6 inches apart, even with each other. Head lowered, shoulders stooped, back bent slightly forward. Standing on flat feet. Leaning and slouching.
Gesture	Speaker used gestures spontaneously.	Absence of hand and arm gestures.
Body Movement	Forward movements for emphasis, sideward movements for transitions as written into speech transcript.	Absence of body movements.

\*For complete validation procedure, see D. F. Gundersen, "Relationships Between Speech Delivery and Speech Effectiveness", Diss. Texas, 1974.

ationalized dichotomously—that is, it appeared in either an “effective” or an “ineffective” condition.

### *Measures of Speech Effectiveness*

The present research employed three sets of criteria which researchers have used to measure speech effectiveness:

1. *Recall-comprehension.* The speech must be understood and remembered for at least a short length of time. Cronkite concludes that speech with no new content to be recalled and comprehended cannot be expected to change attitudes or effect persuasion.<sup>9</sup> Petrie lists nearly one hundred studies which have used recall and comprehension as dependent variables.

2. *Attitude change.* Since much speaking is persuasive in nature, a change in attitude as a consequence of the speech can serve as a measure of speech effectiveness. Bettinghaus, Miller and Hewgill,<sup>10</sup> and Sereno and Hawkins<sup>11</sup> provide examples of this sort of dependent measure.

3. *Ethos.* The intrinsic believability of the speaker following the speech is commonly used as a measure of speech effectiveness. Auditors cannot be expected to find a speech credible if the speaker is not perceived as a credible individual. Hildreth argues that auditors tend to equate perceptions of effectiveness with perceptions of high *ethos*.<sup>12</sup>

<sup>9</sup> Gary Cronkite, *Persuasion: Speech and Behavioral Change* (New York: Bobbs-Merrill, 1969), p. 14.

<sup>10</sup> Gerald R. Miller and Murray A. Hewgill, “The Effect of Variations in Nonfluency on Audience Ratings of Source Credibility,” *QJS*, 50 (1964), 36-44.

<sup>11</sup> Kenneth K. Sereno and Gary J. Hawkins, “The Effects of Variations in Speakers’ Nonfluency upon Audience Ratings of Attitude toward the Speech Topic and Speakers’ Credibility,” *SM* 34 (1967), 58-64.

<sup>12</sup> Richard A. Hildreth, “An Experimental Study of Audiences’ Ability to Distinguish between Sincere and Insincere Speakers,” Diss. Univ. of Southern California, 1953, pp. 71-72.

Andersen and Clevenger list more than twenty studies using *ethos* as dependent measures.<sup>13</sup>

In the present research, actual measures of speech effectiveness were: 1. a test of the recall and comprehension of the main points and important ideas of the speech. Eight such items were administered to thirty-two subjects who had heard the stimulus speech. The six items which proved to be most effective at high-low quartile discrimination were used to measure recall and comprehension. 2. A pre-posttest attitude shift questionnaire. A number of such items were pretested with thirty-two subjects, and the five which provided significant variation from pretest to posttest were employed in data gathering to represent attitude change. 3. A Whitehead source credibility scale providing factor scores indicative of trustworthiness, professionalism, objectivity and dynamism.<sup>14</sup> A general effectiveness question was added to the final questionnaire. This was a five-point item scaled from “very effective” to “very ineffective.”

### *Materials and Procedures*

A persuasive speech concerned with standards for antipollution devices on American automobiles was written as stimulus material for this research. This speech was video-taped in each of ten experimental conditions by the same speaker in the same environment. The ten experimental conditions were: 1. All variables effective condition (EF): all six variables were in the effective state; 2. all variables ineffective condition (IN): all six variables were in ineffective states; 3. verbal ineffective condition (VB): the verbal variables—

<sup>13</sup> Kenneth Andersen and Theodore Clevenger, Jr., “A Summary of Experimental Research in Ethos,” *SM*, 30 (1963), 59-78.

<sup>14</sup> Jack L. Whitehead, Jr., “Factors of Source Credibility,” *QJS*, 54 (1968), 59-63.

volume, voice quality, and rate—were varied together to form a verbal ineffective condition; 4. nonverbal ineffective condition (NV): the nonverbal variables—posture, gesture, and body movement—were varied as a group to form a nonverbal ineffective condition; 5. volume ineffective condition (VO): only the volume variable was in the ineffective state (this condition was the same as the EF condition but had the audio track attenuated 20 db); 6. rate ineffective condition (RA): only the rate variable was in the ineffective condition; 7. voice quality ineffective condition (VQ): only the voice quality variable was in the ineffective condition; 8. posture ineffective condition (PO): only the posture variable was in the ineffective state; 9. gesture ineffective condition; 10. body movement ineffective condition (BM): only the body movement variable was in the ineffective condition.

The video-tape containing these ten conditions of the stimulus speech was presented to nine raters for validity testing. These raters were given written instructions concerning the behaviors they were to rate. In addition to the six independent variables, three distractor items were included in their rating forms: facial expression, appearance and grooming, and eye contact. The raters viewed all ten speeches, completing a rating form after each. Inspection of these forms indicated that each experimental speech was as it was expected to be; that is, only the variables under study were found lacking in each speech tape.

Interestingly, pretests revealed that one distractor variable which remained static in all conditions in the experimental tapes (appearance and grooming) was perceived as having deteriorated in the ineffective condition. This suggests a gestalt phenomenon in speech

rating; when a number of variables are perceived as ineffective, it seems a reverse halo effect may cause perception of ineffectiveness in other variables.

The experimental questionnaires contained an instruction sheet, the attitude pretest, the attitude posttest, the recall-comprehension questionnaire, the ethos scale, and a sheet containing demographic information with the "general effectiveness" question. In all cases the first two pages were the instructions and the attitude pretest, and the last page was the demographic and general effectiveness page. All six orders of the intermediate three pages were used in approximately equal numbers in each experimental condition.

### *Subjects*

Subjects were 229 undergraduates enrolled in courses in the School of Communication of the University of Texas at Austin during the spring of 1974; 68 were freshmen, 79 were sophomores, 52 juniors, and 30 seniors. One hundred and seventy-nine were female and 50 were male. The average age of these subjects fell between 20 and 21, with a range from 18 to 31. There were 123 education majors, 34 communication majors, and 28 business majors. The remainder were scattered among 18 other majors.

The subjects were met by the experimenter in groups of twenty or fewer. Identical video-tape and monitoring equipment was used in all sessions. After the experimental pamphlets were passed out, the subjects were asked to participate in the construction of a standardized speech rating sheet, to view a standard student speech on video-tape and to respond to it on various rating forms.

Subjects first read the instruction sheet and filled out the attitude pretest

page. Then they were seated near the monitor at a distance of ten to fifteen feet to view the stimulus speech. Each group of subjects viewed only one condition of the speech. In all experimental conditions the volume of the monitor was set in the same position. After viewing the speech, subjects filled out the remainder of their questionnaire—containing the recall-comprehension sheet, the posttest attitude questionnaire, and the ethos form, as well as the demographic and general effectiveness items. These experimental sessions typically took from twelve to fifteen minutes. No experimental cell contained fewer than twenty-two nor more than twenty-five subjects.

#### Results

The data were analyzed by one-way analysis of variance. When such analyses indicated significance ( $p < .05$ ), pairwise comparisons were made using Dunnett's test<sup>15</sup> to indicate significance of differences between the total effective condition and the other nine conditions.

The findings, presented by dependent variable follow. *Attitude change*: none of the delivery behaviors exhibited differentially in the experimental stimulus tapes seemed to cause a significant difference in the amount of attitude shift by the subjects. *Recall and comprehension*: subjects in all conditions seemed able to recall the information with approximately the same degree of accuracy. *Ethos—trustworthiness*: this factor of the ethos scale exhibited no significant differences in any experimental condition. *Ethos—professionalism*: this factor of the ethos scale exhibited no significant differences. *Ethos—dynamism*: significant differences among the conditions emerged for the dynamism

factor of ethos ( $p < .001$ ), and Dunnett's test was applied to discover which experimental groups differed significantly from the all-variables-effective group (EF). Table II summarizes these findings.

TABLE II

EXPERIMENTAL CONDITIONS DIFFERING FROM THE EFFECTIVE CONDITION ON THE DYNAMISM FACTOR	
Significant Difference ( $p < .05$ )	No Significant Difference ( $p > .05$ )
IN	PO
NV	VO
VB	VQ
GE	
BM	
RA	

*Ethos—objectivity*: this factor of the ethos scale did not differ significantly among experimental conditions. In sum, the independent variables (delivery behaviors) affected only one of the dependent variables (effectiveness measures) in any significant manner. Several *post hoc* analyses were undertaken as attempts to account for or validate these findings.

#### Post hoc Analyses

Since the subject pool in this research included students from several lower division classes and several majors, it seemed reasonable to suggest that heterogeneity among groups led to the exceedingly large error terms in the analysis of variance. To test this possibility, the data were sorted to isolate the largest single homogeneous group of subjects available, those enrolled in a course in speech for prospective teachers. The results of this analysis were in accord with the original analysis of variance. Only ethos-dynamism emerged salient. This more homogeneous population reacted in much the same fashion as the total subject pool.

To test further the possibility that

<sup>15</sup> C. Mitchell Dayton, *The Design of Educational Experiments* (New York: McGraw-Hill, 1970), p. 49.

differences among the subjects might account for significant portions of the variance in the original analysis, an additional analysis was performed to look for differences between results obtained from education majors, communication majors, and business majors. The analysis indicated that on recall and comprehension (across all independent variables) education majors scored significantly lower ( $p < .05$ ) than did the business and communication majors. No significant differences emerged which could be related to the independent variables. Therefore, it was concluded that within-group variance in the original experiment had not been sufficient to account for the absence of significant findings.

In all previous twelve conditions the content, style and organization (factors of speech composition) of the speech were held constant. Common sense (verified by a follow-up study with twenty-one subjects) indicated that the speech content and composition in stimulus materials was significantly persuasive by itself. Perhaps the content and composition of the speech were strong enough to mask differences which delivery might otherwise produce.

## STUDY 2

Certainly the interaction of delivery and composition is not a novel research topic. McCroskey<sup>16</sup> and Ware and Williams<sup>17</sup> suggest interactions among content, delivery, and ethos. Such lines of reasoning indicate the possibility that since the speech used in this research

was highly persuasive by itself, it may have been primarily the speech's content and composition to which the subjects were reacting, rather than the delivery variables.

## Method

Using procedures derived from those of McCroskey,<sup>18</sup> the stimulus speech was rewritten. Statements attributed to authority were changed to assertions, and appeals to emotions were largely eliminated. In addition, the organization of the speech was revised to yield less obvious connections from point to point. All items required by the recall-comprehension measure were precisely reproduced in the new "ineffective content" speech.

Typescripts of the original and the "ineffective" composition were given to five trained raters, in random order, using two randomly ordered rating instruments containing the following scales on a 7-point continuum: well-badly composed, stylish-lacks style, convincing-unconvincing, good-bad, effective-ineffective, and an "A" speech-an "F" speech. On all six scales the new speech rated at least three points lower than the original speech.

This "ineffective composition" speech was video-taped in the same setting, using the same speaker, clothing, and equipment as the original ten speeches. This speech was delivered in only two conditions, comparable to the "total effective" and "total ineffective" conditions of the initial study.

These two stimuli were presented to five trained raters and were rated using the same validation instruments used in the initial study. The validity data corresponded closely to the data of the original speeches, indicating that the delivery behaviors were similarly executed

<sup>16</sup> James C. McCroskey, "A Summary of Experimental Research on the Effects of Evidence in Persuasive Communication," *QJS*, 55 (1969), 169-76.

<sup>17</sup> John E. Ware, Jr. and R. G. Williams, "An Experimental Study of the 'Doctor Fox' Effect," paper presented at the 18th International Congress on Applied Psychology, Montreal, Quebec, July 1974.

<sup>18</sup> McCroskey, "Effects of Evidence."

in these congruent delivery conditions. The similar delivery ratings, coupled with the dissimilar content ratings, provided sufficient validity data to posit that this second set of stimuli was comparable to the effective conditions in delivery, but not in content.

The effective-delivery-ineffective-composition condition was shown to thirty-two subjects, the ineffective-delivery-ineffective composition condition to nineteen. Both groups viewed the stimuli under the same conditions as outlined in the initial study. Data from these two sessions were combined with data from the original "effective" and "ineffective" conditions and were analyzed using two-way analysis of variance.

### Results

Attitude change was sensitive to composition differences ( $p < .01$ ). Recall-comprehension was not affected by composition, delivery, nor their interaction. Ethos-trust-worthiness was significantly affected by composition ( $p < .01$ ), delivery ( $p < .05$ ), and their interaction ( $p < .05$ ). Ethos-professionalism was significantly affected by composition ( $p < .01$ ), delivery ( $p < .01$ ), with a trend toward an interaction ( $p = .10$ ). Ethos-dynamism proved sensitive to differences in both delivery ( $p < .05$ ) and composition ( $p < .01$ ) with no interaction. Ethos-objectivity was sensitive to composition ( $p < .01$ ) but not to delivery with no interaction evident. The general effectiveness question was sensitive to composition ( $p < .01$ ) and delivery ( $p < .05$ ) with a trend toward an interaction ( $p = .10$ ).

These data indicated that attitude change was sensitive to manipulations in the composition of the stimuli. Recall and comprehension were not affected by any of the variables manipulated in the post hoc study. Ethos-trustworthiness and professionalism, as well as the

general effectiveness item, varied according to composition and according to delivery in the ineffective content condition. Ethos-dynamism varied in both content conditions and in both delivery conditions. Ethos-objectivity was primarily affected by composition.

### DISCUSSION

All the delivery variables operationalized in this speech are highly regarded by writers of texts in platform speaking. The present study suggests that these aspects of delivery are secondary to composition in ability to bring about measurable speech effectiveness. Effective content served in this study, to mask the effects of ineffective delivery in all time conditions.

This is not necessarily an argument for discontinuing the teaching of speech delivery, for in the ineffective-speech-composition. Since speech compositionally made an ineffectively composed speech nearly as effective as speeches of good composition. Since speech composition will be mastered in various degrees by students and since the persuasiveness of speech messages cannot often be pretested on the target audience, perhaps speech delivery should be taught to students as an additional, but secondary, method of achieving gains in effectiveness. This research supports the teaching of speech delivery *only* as a secondary method of achieving speech effectiveness; teaching composition seems significantly more effective.

The present research leaves a number of questions unanswered. For instance, what are the effects of the variables studied here on longer-term attitude change? In this same vein, since delivery behaviors have significant effects on most measures of speech effectiveness only in low composition conditions, do changes in a single variable (as in the

initial phase of this research) produce similar consequences?

The recall-and-comprehension variable leaves the question of the effect of the variables used here on long-term recall. It is possible that both delivery and composition affect recall differently over a period of time. The ability of listeners to integrate and apply knowledge gained would also be a relevant subject for study.

The findings of this research supplement those of McCroskey<sup>19</sup> in suggesting interactions of delivery and composition with ethos. These interactions could be taken to exist for the total of

ethos, or for each of its factors; the interactions could be varied by pretesting both the speech and the speaker for the various factors of ethos and discovering the interactions for delivery and composition variables crossed with the factors of ethos.

There seems to be a place in this type of research for the construction of regression models to determine which of the various delivery-composition-ethos variables best predict general effectiveness. The present research indicates that the single general effectiveness five-point scale is a good measure of effectiveness as measured by any other dependent variables. Future research might well investigate these implications.

<sup>19</sup> *Ibid.*



