A study examined the relative effectiveness of written feedback on student speeches. Subjects, 114 students from a large, midwestern university enrolled in 10 sections of the basic communication course, supplied a photocopy of a speech evaluation form which included written comments as well as an overall speech grade for the third speech of the semester. A total of 2,933 comments on 114 evaluation forms were coded for 4 dimensions. Results indicated that: (1) positive comments were more prevalent than negative remarks; (2) there was a direct relationship between positive written feedback and grades; and (3) evaluation forms contained significantly more written comments directed toward content (63%) than those directed toward delivery (28%), outlines, bibliographies and time constraints (6%), or general comments (3%); (4) evaluations tended to have significantly more multi-word comments than single-word comments; and (5) there was no significant difference regarding gender and observation values. One table presenting operational definitions and 2 tables of data are included. (Contains 27 references.) (RS)
Evaluation of Student Speeches:
A Content Analysis of Written Feedback
Karla Kay Jensen, Elizabeth R. Lamoureux, Jeanette E. Brossmann, and Diana Prentice-Carlin
Kansas University
Department of Communication Studies

Running head: EVALUATION OF STUDENT SPEECHES
March 8, 1993

*An earlier version of this paper was presented at the Annual meeting of the Speech Communication Association, Chicago, IL, October 29 - November 1, 1992.

*We would like to acknowledge Dr. Mary Lee Hummert for helping generate the idea for this study as well as her recommendations and reading of various versions of the manuscript.
Evaluation of Student Speeches: A Content Analysis of Written Feedback

At its core, teaching appears to be a simple endeavor. It seems to involve little more than providing information to students and evaluating how well they master it. Of course, a meaningful education involves much more, including the skill with which teachers reveal their evaluations to students. One dimension of teaching effectiveness is encapsulated in Holtzman's (1960) question: "What can I write (or say or do) that will result in this student's improving his or her communicative ability?" (p. 1).

As instructors evaluating student performances in the communication basic course, advanced persuasive speaking, or speech writing we find ourselves repeating Holtzman's query on a daily basis. Every time a student presents an oral presentation, we feel compelled to provide a written evaluation which contains not only a summative discussion to assess strengths and weaknesses, but a formative dimension to shape future performances. Staton-Spicer & Wulff's (1984) content analysis of articles examining the effectiveness of written feedback seems to suggest that there is little direction for those of use seeking to improve both levels of evaluation. Of the 186 SCA journal articles dedicated to communication education or instructional communication written between 1974 and 1982, only seven directly addressed the criticism of students' oral performances. Of those seven, only one examined instructors' written feedback. It would appear that we know surprisingly little about one of our most
frequently used methods of providing instructional communication to our students.

The current study is an initial attempt to garner information about the relative effectiveness of written feedback of student speeches. We explored such issues as the types of comments made by instructors, the areas upon which those remarks focus, the proportions of positive and negative observations, the relationship between positive statements and grades, and the effects of student gender. Our rationale is simple: we believe that Cathcart's (1966) reminder that learning cannot take place without criticism is of little value if we do not know whether our criticism actually facilitates our students' learning.

REVIEW OF LITERATURE

In comparison to other forms of feedback, written evaluations appear uniquely valuable. While audio and video tape have become useful supplements for verifying and reinforcing comments (Bradley, 1970; Deihl, Myles & Larson, 1970; Fuller, Veldman & Richek, 1966; Garrison, 1984; Hirschfeld, 1968), the difficulties associated with recording feedback have generally limited their use to enhancing personal consultations between students and instructors (Deihl et al., 1970; McCroskey & Lashbrook, 1970). Further, while verbal feedback is undeniably valuable, factors such as class size and time constraints restrict its use. Thus, for students who may not remember the full range of oral comments, written criticism has the advantage of both immediate and future reference.
In addition to this apparent advantage, written feedback also appears to motivate student achievement and improve speech performance (Book & Simmons, 1980). However, its successes would seem to be tied to its effective use. What constitutes effective use of written feedback of student speeches is currently unknown. First, most of the feedback literature focuses on oral, not written criticism (Book, 1983; Bostrom, 1963; Dedmon, 1967; Roubicet, 1990; Staton-Spicer & Wulff, 1984). Second, the few studies on written feedback offer limited information on actual classroom practices (Miller, 1964; Palmerton, 1986; Sprague, 1971; Young, 1974). Finally, most of these studies were conducted in the early 1970s, suggesting a need to confirm their contemporary applicability. Friedrich (1981) has suggested a variety of changes in the classroom which require teacher adaptation, including changes in student demographics, budget constraints, technological advances, and greater accountability for quality instruction. Given the current lack of information about written feedback in public speaking classes, this study explores how instructors employ written feedback through our initial research question:

RQ1: What forms of written comments do basic course instructors use in evaluating speeches?

In addition to the types of remarks instructors make, there is a growing concern about gender variations in instructional evaluations. Research in a variety of disciplines has shown that the biological gender of the instructor is influential in the evaluation process (Sandler, 1986). Although Sprague (1971)
hypothesized there would be no significant difference in critique patterns in terms of gender of instructor, gender of student, or a combination of both, her findings contradicted her hypothesis. Critique patterns differed significantly for male and female instructors, male and female students, and for various combinations of the two. Because previous research has largely focused on instructor gender, the current study expands the scope by examining the effects of student gender. Because Sprague's (1971) research is now more than a generation old, and given the renewed societal interest in gender awareness (Sandler, 1986), we believe this issue is especially pertinent at this time. Thus we are seeking to answer the question:

RQ2: What type of biological gender differences exist in written speech evaluations?

In addition to questions of form and gender, content valence is an important issue in written evaluations. Instructors must continually balance the competing needs of exposing weaknesses in performance without unduly damaging student confidence. Previous research suggests that speech evaluators use at least as many, if not more, positive than negative comments (Book & Simmons, 1980; Young, 1974). Our first hypothesis attempts to support those findings.

H1: Written speech evaluations will contain more positive than negative comments.

The valence question also has implications for students' grades, although the relationship is not clear. Sprague (1971) found that critique patterns differed significantly for the
various letter grade groups. She found that more positive comments were written for "A" speeches than for "B" speeches, for "A" and "B" speeches than for "C" speeches, and for "A", "B" and "C" speeches than for "D" speeches. Although other researchers have incorporated student grades into their evaluation research (see Hammer, 1972; Holtzman, 1960; Stewart & White, 1976; Vogel, 1973), the results are not uniform. For example, the findings of Book & Simmons (1980) and Young (1974) conflict with Sprague's (1971) reported relationship between speech grade and comment valence. While we recognize that it is sometimes important to provide more encouragement to struggling students, we believe that Sprague's results more accurately reflect the common relationship between evaluations and grades as stated in our second hypothesis.

**H2:** There is a positive relationship between positive written feedback and grades.

Finally, we realize that the lack of consensus on the relative importance of content and delivery in public speaking is at least 2,000 years old. While we take no position on this debate, we do assume that critics are likely to spend more time discussing elements they find to be the most important. Further, we believe that students are likely to view the proportion between content and delivery comments as indicative of their relative importance. Consequently, regardless of one's personal views on this issue, the ratio of content to delivery remarks would appear important. In her own study, Sprague found that 73% of all comments dealt with content, while 27% focused on delivery.
Although we have identified two additional categories of comments (outline, bibliography & time, and general comments), we expect to find more content comments than any other type. Thus we hypothesize:

H3: There will be significantly more content comments than other types of remarks.

The remainder of this essay outlines the methodology employed in this study, reports the results, and discusses the findings and their implications for written evaluations of students oral performances.

METHOD

Subjects and Procedure

One hundred and fourteen students from a large, midwestern university were drawn from ten sections of the basic communication course. Students ranged in age from 18 to 21, with a mean age of 19. Participants supplied a photocopy of a speech evaluation form which included written comments as well as an overall speech grade. Evaluation forms were gathered from 48 males and 66 females, all of whom read and signed consent statements ensuring their confidentiality.

The evaluations provided a sampling of comments from ten course instructors as well as from speeches receiving grades across the entire grading scale. Eleven to twelve forms were gathered from students in each of the ten participating sections. These forms were collected from the third speech of the semester, an informative presentation. The third speech was selected because of the potential for atypical comments in the first or
final speech evaluations. The first speech is a "trial run" for both the student and the instructor, since both are assessing the student's capabilities and potential. These first comments may be exceedingly encouraging or general and therefore unrepresentative. As Sprague (1971) notes, the final speech may be equally unrepresentative, since it is often the "last best effort," and may result in proportionately more positive comments than other speeches. Therefore, a middle speech appeared to be an appropriate selection for our research since a certain level of mastery is expected, yet comments also focus on future goals and improvement.

Instrumentation

Our coding scheme was adapted from the work of Sprague (1971) who proposed a category system based on four dichotomies: 1) content-delivery, 2) atomistic (specific) - holistic (general), 3) personal-impersonal, and 4) positive - negative. Sprague's categories have been utilized in earlier speech evaluation research: evaluation from instructors (Vogel, 1973); evaluation from peers (Book & Simmons, 1980); and evaluation of student preferences for written comments (Young, 1974).

Sprague's (1971) findings revealed that instructors' written evaluations contained significantly more atomistic than holistic comments. Additionally, Young (1974) found that students identified specific comments as significantly more helpful than generalized statements. In a similar study by Book & Simmons (1980), atomistic comments by peers were also consistently seen as more helpful than holistic comments. In addition to the
consensus regarding specific comments, the research consistently reveals that impersonal comments are perceived as more helpful than personal comments (Book & Simmons, 1980; Sprague, 1971; Young, 1974). However, the personal-impersonal category presents coding problems because it is difficult for coders to know the relationship between the student and instructor.

Since the research shows that instructors are already providing the types of comments students find most helpful in the aforementioned areas, we chose to focus on other important issues surrounding the written critique. Because the literature evidences a continuing controversy in the categories of positive-negative and content-delivery comments (Book & Simmons, 1980; Sprague, 1971; Young, 1974) we chose to evaluate these two areas. However, there are several rationales for investigating other issues concerning the written speech critique.

First, since observation indicates that instructors provide other types of comments, including general remarks and observations about outlines, bibliographies and time, we decided to expand Sprague's (1971) content-delivery dichotomy to provide a more representative typology of speech comments. Second, written evaluation comments contain a variety of forms and lengths. Since the goal of this study is to enhance understanding of written evaluations, we included an analysis of both of these dimensions. Given the preceding considerations, this study examines comment type, valence, length and form.
Comment type, operationalized as content, delivery, general and OBT (outline, bibliography, or time), refers to the type of information provided in each comment. Content included remarks about ideas, reasoning, supporting material, organization or language. Delivery involved comments concerning the physical or vocal elements of communication. Outline, bibliography, or time (OBT) references are self-explanatory. Finally, general comments included references to the speech as a whole or categories other than content, delivery or OBT.

Comment valence involved analyzing the positive or negative dimensions of a comment. Positive comments included those which complimented or expressed approval of the presentation. In contrast, negative comments expressed disapproval or made suggestions for improvement. All comments were found to have either a positive or negative dimension.

Content length simply differentiated between one-word comments (single-word) and multiple-word comments (multi-word). The final comment dimension under investigation involves comment form. This dichotomy involves questions (interrogative comments) and statements (declarative comments).

Coding Procedures

The unit of analysis for this study was the topical phrase. A topical phrase is a comment that can stand alone (a word, phrase or clause that relates to one topic in the evaluation). Content analysis was used because it operates directly upon
transcripts of human communication, yielding unobtrusive measures, and utilizing both qualitative and quantitative modes of analysis (Holsti, 1969; Krippendorf, 1980; Weber, 1985). The comments were content analyzed for comment type, valence, length and form. Thus, a statement such as "1) Good introduction, 2) but you needed more eye contact, 3) and a clearer thesis," would be coded as three, separate constructs: 1) content / positive / multi-word / statement; 2) delivery / negative / multi-word / statement; and 3) content / negative / multi-word / statement.

In the current study, the number of comments per evaluation ranged from 5 to 54 with a mean of 24 comments. The mode was also 24.

Three trained coders were familiarized with the instrumentation and the purpose of the study. A total of 2933 comments contained on 114 evaluations were coded for all four dimensions. Intercoder reliability, calculated according to Holsti's (1969) formula, was .91.

RESULTS

All three hypotheses were supported. In support of the first hypothesis, positive comments were more prevalent than negative remarks. Positive observations comprised 52% of the sample, with negative remarks accounting for 48% ($X^2 (1) = 3.90$, $p < .05$).

The relationship between comment valence and overall grade was also supported, although small cell sizes required that speeches earning grades below C be excluded from the data. A
one-way analysis of variance revealed a significant difference between grade and the number of positive comments ($F(2, 106) = 8.8117, p < .01$). Post-test analyses revealed that all three groups were significantly different from each other ($p < .05$). An analysis of means indicated that A speeches received significantly more positive comments than B speeches, and B speeches received significantly more positive comments than C speeches (see Table 2).

The final hypothesis received the strongest support. Evaluation forms contained significantly more written comments directed toward content (63%) than those directed toward delivery (28%), outlines, bibliographies and time constraints (6%), or general comments (3%) ($X^2(3) = 2704, p < .001$).

The results of the third hypothesis also shed light on the first research question, which concerned the nature of written evaluations. In addition to the distinctions between content, delivery, OBT and general comments, the evaluations also contained significantly more multi-word (89%) than single-word (11%) comments ($X^2(4) = 5.98, p < .05$) (see table 3).

The last research question involved the role of student gender on the nature of written evaluations. This study failed
to find a significant difference regarding gender and observation valence ($X^2 (1) = 1.45, p < .25$). While 52.9% of the comments directed toward females were positive, 50.6% of the comments directed toward males were also positive. Significant gender differences were found concerning the types of comments provided. Females received a greater percentage of content-delivery comments than males ($X^2 (1) = 3.70, p < .05$). While 33.4% of the content-delivery comments directed towards females concerned delivery, 28.9% of content-delivery comments directed toward males involved delivery.

**DISCUSSION**

The results illuminate a variety of aspects concerning written speech evaluations. The first hypothesis, indicating that positive comments will be more common than negative comments, was supported. This result is consistent with current research that encourages the use of positive feedback. Specifically, comments that are encouraging and personalized tend to be perceived by the students as the most effective (Stewart & White, 1976). Additionally, Bostrom (1963) found that positive reinforcement from the instructor tended to yield more positive student attitudes toward the speaking experience. In contrast, negative comments should also be included on evaluations. First, learning theory indicates that allowing undesirable behavior to continue without comment reinforces that behavior (Young, 1974). Second, some scholars argue that people do not necessarily desire feedback, and that excessive praise may have damaging effects,
resulting in a lack of further motivation and/or overconfidence (Farson, 1963).

The strong support for the second hypothesis, regarding a direct relationship between positive comments and grade, reinforces Sprague's (1971) findings. However, it is important to note that due to a lack of D and F speeches it is impossible to say if this relationship is consistent across the grading spectrum. In fact, the only D speech included 18 positive comments, which is more than the means for either B or C speeches. Future research may want to focus on the relationship of positive comments to all possible grades.

With regard to speech grade, it is not the comments alone, but the comments in conjunction with a letter grade which would be more likely to improve student performance (Stewart & White, 1976). This evidence is particularly valid where individually tailored comments take into consideration the student's grade expectation. Because our experience shows that students are generally preoccupied with grades, this could be a topic worth additional attention.

The results of this study also suggest that speech evaluations tend to have significantly more content comments than any other type of comment. The strong support for the third hypothesis mirrors the emphasis placed on content at the university under investigation. These findings are also consistent with Sprague's (1971) research. While Sprague did not include a general or OBT category, her research reported that 73% of the content-delivery comments concerned speech content.
Considering the role of a basic public speaking course, where developing a well-organized, well-researched speech is the foundation, the finding that more comments are devoted to content is encouraging. For most novice speakers, delivery will develop over time, especially if students are confident in their abilities to create a solid speech. Young (1974) discovered that from students' perspective, delivery is a reflection of their total being. While the content of a speech may also reflect personality, students may feel that it is easier to improve content without changing one's self. Therefore, students may be more sensitive to delivery criticism. Instructors should also be aware that the evaluation form may influence the emphasis given to content and delivery. This is particularly true for the standardized evaluation form used at the university in this study, where the format encourages instructors to write more content than delivery comments.

This study also found that evaluations tended to have significantly more multi-word comments than single-word comments. This finding may indicate an awareness on the part of instructors regarding the need to clarify feedback by providing more information concerning a given comment. We welcomed this finding since our own teaching experiences have demonstrated that students prefer detailed, multi-word phrases over single-word remarks. Future research might focus on how students perceive the length of feedback.

Another interesting finding surrounds the influence of speaker gender on written evaluations. This study failed to find
a significant difference of comment valence and gender. These findings diverge from gender differences found in related areas. For example, Book & Simmons (1980) found that women received more than twice as many positive comments as men. Additionally, Sandler's (1986) gender research in academic settings revealed that professors, regardless of biological gender, tended to create a more positive learning environment for male students by giving them a greater share of classroom attention. Although the past decade has produced disagreement among scholars about gender issues in the classroom, the results of the current study may reflect trends in the attempt to achieve greater equity concerning gender across a variety of contexts, including higher education.

While many of the results of the current study are statistically significant, we hope that they can also be educationally significant. Future research could examine students' perceptions of the written evaluation process. Leauby and Atkinson (1989) challenge teachers to create an atmosphere which enables students to motivate themselves to achieve their potential. Students, particularly at the university level, are valued parties in this endeavor. Efforts to corroborate these results should consider student perspectives.

As educators we have the responsibility to investigate our own teaching practices for evaluation and improvement. Criticism of student speeches must be carefully planned and executed if speech instructors are to modify and improve the speaking skills of students. Assessing written comments of actual students'
speeches can aid in this process. Because of its practicality, this naturalistic and descriptive study should produce a relevant foundation for the continued investigation of written feedback on student speeches, thus providing both interest and value to instructors of the basic course as well as our students.
References


Sandler, B.R. (1986). *The campus climate revisited: Chilly for women faculty, administrators and graduate students*. (Publication of the project on the status and education of women association of American colleges [Washington, D.C.]).


Table 1
Operational Definitions

Comment Type:

1) Content – Any comment dealing with ideas, reasoning, supporting material, organization, or language.
Examples: Appropriate selection and use of support materials; The main points were difficult to distinguish.

2) Delivery – Any comment dealing with the physical and vocal elements of communication such as rate, volume, fillers, inflection, eye contact, gestures, posture, poise, articulation, pronunciation, dynamism, sincerity or confidence.
Examples: Excellent eye contact; Work on articulation.

3) Outline, Bibliography, Time (OPT) – Any comment which addresses the outline bibliography or time constraints.
Examples: Outline has nice structure; Bibliography needs to be alphabetized.

4) General – Any comment which views the speech as a whole, or addresses topics other than those already categorized.
Examples: Your efforts are appreciated; More preparation would have resulted in a stronger speech.

Comment Valence:

1) Positive – Any comment which compliments or expresses approval of the speaker or the presentation.
Examples: Professional stance; Great enthusiasm; Original topic.

2) Negative – Any comment which expresses disapproval or makes suggestion for improvement.
Examples: Work for greater vocal variety; Use more transitions so audience is able to follow your speech.

Comment Length:

1) Single-word – Any comment which is limited to one word.
Examples: Work for greater vocal variety; Use more transitions so audience is able to follow your speech.

2) Multi-word – Any comment which uses two or more words.
Examples: Solid credibility; Incorporate visual aids earlier.

Comment Form:

1) Question – Any comment which asks a question.
Examples: What? Date? Where did you find your supporting material?

2) Statement – Any comment which is declarative.
Example: Need to analyze your audience; What a memorable final thought!
Table 2
Mean of Positive Comments by Speech Grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>19.37</td>
<td>13.67</td>
<td>6.00</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>n</td>
<td>(27)</td>
<td>(59)</td>
<td>(23)</td>
<td>(1)</td>
<td>(1)</td>
</tr>
</tbody>
</table>
Table 3

Frequencies of Categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of Constructs</th>
<th>Percent of Total Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>1849</td>
<td>63%</td>
</tr>
<tr>
<td>Delivery</td>
<td>822</td>
<td>28%</td>
</tr>
<tr>
<td>General</td>
<td>90</td>
<td>3%</td>
</tr>
<tr>
<td>(OBT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outline, Bibliography, Time</td>
<td>172</td>
<td>6%</td>
</tr>
<tr>
<td>Positive</td>
<td>1520</td>
<td>52%</td>
</tr>
<tr>
<td>Negative</td>
<td>1413</td>
<td>48%</td>
</tr>
<tr>
<td>One-word</td>
<td>318</td>
<td>11%</td>
</tr>
<tr>
<td>Multi-word</td>
<td>2615</td>
<td>89%</td>
</tr>
<tr>
<td>Question</td>
<td>98</td>
<td>3%</td>
</tr>
<tr>
<td>Statement</td>
<td>2835</td>
<td>97%</td>
</tr>
</tbody>
</table>