


2002

## An Examination of the Speech Evaluation Process: Does the Evaluation Instrument and/or Evaluator's Experience Matter?

Karen Anderson  
*University of Kansas*

Karla Kay Jensen  
*Nebraska Wesleyan University*

Follow this and additional works at: <http://ecommons.udayton.edu/bcca>

 Part of the [Higher Education Commons](#), [Interpersonal and Small Group Communication Commons](#), [Mass Communication Commons](#), [Other Communication Commons](#), and the [Speech and Rhetorical Studies Commons](#)

---

### Recommended Citation

Anderson, Karen and Jensen, Karla Kay (2002) "An Examination of the Speech Evaluation Process: Does the Evaluation Instrument and/or Evaluator's Experience Matter?," *Basic Communication Course Annual*: Vol. 14 , Article 10.  
Available at: <http://ecommons.udayton.edu/bcca/vol14/iss1/10>

This Article is brought to you for free and open access by the Department of Communication at eCommons. It has been accepted for inclusion in Basic Communication Course Annual by an authorized editor of eCommons. For more information, please contact [frice1@udayton.edu](mailto:frice1@udayton.edu), [mschlangen1@udayton.edu](mailto:mschlangen1@udayton.edu).

## **An Examination of the Speech Evaluation Process: Does the Evaluation Instrument and/or Evaluator's Experience Matter?**

---

*Karen Anderson  
Karla Kay Jensen*

Characterizing the public speaking course as “bedrock of the undergraduate curriculum” (p. 75), Lucas (1999) recognizes that objectively assessing the quality of a student’s work can be one of the most challenging tasks for those teaching this course. Consequently, the speech evaluation process, instrument design and its use are critical to those with a vested interest in improving the basic course. By using evaluation instruments commonly found in the public speaking classroom, we attempt to determine whether the instrument or the raters’ level of experience influences the grading process. Second, through surveys and open-ended questionnaires, we examine evaluators’ perceptions and use of these evaluation forms.

### **THE SPEECH EVALUATION PROCESS**

Many public speaking texts, instructor manuals, and other guides contain speech evaluation instruments in an effort to establish criteria by which speeches will be evaluated. Rubin’s (1999) suggestion that the basic

principle of setting criteria before evaluating has guided our discipline for the last 70 years. Yet even with pre-determined criteria, raters can give biased evaluations. For instance, leniency errors can (Bock, 1970) occur when the evaluator is either too easy or too harsh on all speakers. Halo effect errors (Bock, 1974) can occur when the evaluator is either too easy or too harsh on a particular speaker. Both of these errors can occur when evaluators are aware that the student will see the results (Bohn & Bohn, 1985). Additionally, Stiggins, Backlund and Bridgeford (1985) recognize that lack of training, the evaluator's culture, and even the perceived anxiety of one's students can lead to increased rater bias.

Other studies illustrated that rater training (Bohn & Bohn, 1985, Bowers, 1964; Gunderson, 1978; Miller, 1964), experience (Clevenger, 1963), or the combination of the two (Miller, 1964) improved the evaluation process. For instance, Bowers found that when a group of instructors were trained, the variations among their grades was much lower than a group that received no training. Bohn and Bohn's study "graphically demonstrated . . . not only will training reduce rating error, it will also help to improve student speaking performance" (p. 350). Although there has not been any research on the differences in grading good speeches versus poor speeches, Roubicek's (1990) work examined feedback given by novice and experienced instructors. His study found that there were no considerable differences in how each group offered feedback to their students.

In general, studies conclude that the evaluation instruments can and do affect the judgment of the rater (Becker, 1962; Brooks, 1957; Clevenger, 1964; DiSalvo

& Bochner, 1972). Thus, in tandem with researching instructor experience, it is imperative that evaluation design be evaluated. Research suggests that evaluation instruments can be more reliable and valid if they are simple and balanced in terms of content and delivery, and a speaker's overall effectiveness (Holtzman, 1960; Young, 1974). In addition, the qualities outlined in the evaluation instrument should "be those that are emphasized and taught in the class" (Rubin, 1999, p. 428). Several evaluation instruments have been and are still being developed by instructors and authors, however, these may have not been tested for reliability and validity. In contrast, the *Competent Speaker Evaluation Form* (CSEF), developed and tested in 1990 by Morreale and an SCA/CAT subcommittee is "anchored in the communication literature regarding competent public speaking" (Morreale, Whitney, Zautke, Ellis, McCormick & Whitter, 1992, p.10). This instrument, which has been tested to be reliable and valid, is comprised of eight public speaking competencies including: 1) choosing and narrowing a topic for the audience and occasion; 2) communicating the thesis/specific purpose in a manner appropriate for the audience and occasion; 3) providing appropriate supporting material based on the audience and occasion; 4) using an organizational pattern appropriate to topic, audience, occasion and purpose; 5) using language that is appropriate to the audience and occasion; 6) using vocal variety in rate, pitch and intensity to heighten and maintain interest; 7) using pronunciation, grammar and articulation appropriate to the designated audience; 8) using physical behaviors that support the verbal message. The form scores each competency as unsatisfactory, satisfactory

or excellent and allows room for general comments. Although there has been criticism of the CSEF (Hugenberg & Yoder, 1996), most evaluation instruments located contain the eight CSEF competencies.

Although evaluation forms seem to contain common themes, the directions that accompany evaluation instruments can be varied or non-existent. Brooks (1957) contends that directions should be "precise and complete" (p. 29). In addition, Brooks cites various authors who concur that directions encouraging accuracy rather than speed result in more reliable evaluations. Clevenger's (1964) research discovered that a general evaluation form was less reliable than one that directed the raters to evaluate specific qualities of the speech. DiSalvo and Bochner (1972) found that raters do not always use evaluation forms as they were intended to be used. Specifically, participants overwhelmingly used the items that clustered around the concepts of "language" and "delivery" to evaluate the speech. Items of organization, analysis, and speaker personality were seldom used to determine the speech grade. Further, one quarter of the participants used only one item to grade the speech.

Recently, Carlson and Smith-Howell (1995) continued speech evaluation research by examining the reliability and validity of various instruments used in their own department and/or by the participants in their study. First, the expected rating of two video-taped speeches was determined. Next, the instrument reliability was measured by examining the scores given to the two speeches by three levels of raters (experienced from the speech staff, moderately experienced from the mass media staff, and inexperienced from undergradu-

ate public speaking students) using four evaluation forms. Three of these forms were consistent throughout the project, while one form varied from rater to rater. The researchers concluded that a variety of instruments can be used effectively as long as they account for content and delivery. Carlson and Smith-Howell also maintained that the lack of extensive training did not have a major negative impact on the speech evaluation process.

Carlson and Smith-Howell's (1995) study design has a few potential concerns that should be addressed. First, the 58 participants evaluated each speech twice, using two different evaluation forms. Multiple exposures to a speech could influence the perspective the evaluator has regarding that speech. A larger pool of raters might avoid this problem. Second, not all four evaluation forms were used by participants in all three experience levels. For example, Form D was used only by the experienced participants and Form C was never used by this group. Thus, the claim that experience does not matter is perhaps an overstatement. In order to fully support the claim that experience level did not influence the ratings, participants at all experience levels must use all forms in the design. Additionally, the cells of the sub-groups were drastically out of balance. Specifically, of the 19 moderately experienced participants, only two used Form C on Speech 1 and 2, the experienced participants never used Form C on either speech, and the moderately experienced and inexperienced participants never used Form D on either speech. Although the difficulty of finding participants is recognized, and although statistical procedures can adjust for this factor, more balance among the sub-groups might have yielded different results. Finally, although not a methodological

concern, the Carlson and Smith-Howell study contradicts many of the previously held opinions about the importance of training in order to avoid rater bias (DiSalvo & Bochner, 1972; Rubin, 1999; Stiggins et al., 1985).

In addition to these concerns, none of the previously cited research address *how* evaluators use speech evaluation forms or the directions that accompany them. LaLumia (1993) points out that most evaluation forms cover areas identical or similar to delivery, language, organization and purpose and that, despite forms' similarities, teachers may use the instruments in different ways to "fill the particular needs of their programs" (p. 241). For example, evaluators may accurately follow directions on the form, or they may make the form fit the grade they believe should be assigned to the students. We can investigate the use of evaluation forms further by asking about an evaluator's like or dislike of the form. Answers to these questions are important to both students and teachers since over 50 percent of the final grade in many basic course programs is comprised of oral performance grades (Gibson, Hanna & Huddleson, 1985). If a particular form is being used in a variety of ways it could yield different grades. This obviously has implications on issues of grade inflation or deflation and consistency across a large number of sections of the basic course within a given department.

By reviewing the previously cited literature, most of which is dated, it is clear that additional speech evaluation form research is warranted. Such an investigation may aid pedagogues in the creation of evaluation forms, as well as assist in the training of how to best use them.

In order to examine these concerns, four research questions were explored:

**RQ1:** Does previous speech grading experience affect speech ratings?

**RQ2:** Do raters who have written directions on how to use evaluation forms rate speeches differently than raters who do not have evaluation form directions?

**RQ3:** Do evaluation forms affect speech ratings?

**RQ4:** What are evaluators' opinions of the evaluation forms they use?

## **METHOD**

### ***Subjects***

In order to avoid the problems of multiple exposures that Carlson and Smith-Howell (1995) faced, the researchers had a total subject pool of 112 participants. Forty-seven were men; 65 were women. Participants were categorized in three age groups. Sixty-one evaluators fell into the 18-23 age group; 22 participants were in the 24-29 group; 29 participants were in the 30 and over group. The participants were gathered from a variety of locations. Twenty-seven percent were from the Communication Studies department, seven percent were from the Mass Communication department, seven percent were high school speech teachers, nine percent were forensics students, 38 percent were undergraduate students and 12 percent were from Toastmasters.

The participants were grouped into three categories: experienced, moderately experienced, and inexperienced



raters. The 33 experienced raters had over six months of rating speeches in the classroom or in another venue (i.e., Toastmasters or high school teaching) and specific training in rating speeches. These raters included full-time faculty members at two universities, graduate teaching assistants at one university, high school speech teachers at one high school, and members of the local community. The 31 moderately experienced raters had less than six months experience rating speeches in the classroom or no rating experience in the classroom, but related experience in forensics or broadcasting. These raters included full-time faculty members from mass-media, undergraduate forensics competitors who had judging experience, incoming teaching assistants, and members of the local community. The 48 inexperienced raters had no formal rating experience in a competitive or educational setting and had not taken public speaking or another related course that may have influenced their perception of the rating process. This group, comprised of undergraduate students with no public speaking experience, was chosen in an attempt to answer the Carlson and Smith-Howell (1995) suggestion of using such a group to better understand how novice evaluators grade speeches. In addition, many basic course programs employ Master's candidates who have recently completed BA degrees in a variety of disciplines. Studying inexperienced raters may give us insight into how novice TAs would perform without training.

## ***Evaluation Instruments***

Four speech evaluation forms were selected to evaluate a variety of assessment techniques with minimum duplication. All forms selected were representative versions of common instructor-generated instruments published in instructor manuals and/or used at various universities.<sup>1</sup> First, Form A (see Appendix A), used by a large mid-western university and a large southwestern university, is a 100-point scale, which accounts for content and delivery features. The maximum points for each section includes: introduction, 20; body, 35; conclusion, 15; and delivery 30. Within each section a checklist is provided with numerous criteria. This form includes detailed descriptions of what constitutes an A, B, C, D or F speech, using the standard grading scale of 90-100=A; 80-89=B; 70-79=C; 60-69=D; and 59 and below=F. These instructions were given to half of the raters during the study, while the other half of the raters received no instructions other than the point values that were printed on the form.

Form B (see Appendix B) is a “commonly recommended evaluation form” (Carlson & Smith-Howell, 1995), chosen because it was used in the Carlson and Smith-Howell’s study. This form focuses on five areas of concentration: introduction, organization, development,

---

<sup>1</sup> Similar to the Carlson and Smith-Howell (1995) design, the evaluation instruments used in this study all contained recommended and previously studied components of content and delivery. See Sprague (1971), Jensen and Lamoureux (1995) and Rubin (1999) for summaries of evaluation instruments.

conclusion, and delivery. Each area is rated on a five-point scale (poor, fair, average, good, and excellent). The final grade is determined using the same standard scale as Form A. The instructions for this form are not as detailed as Form A or C and include only basic guidelines for evaluating a speech. Half of the raters received instructions and half were not provided with any instructions during the study. The instructions that were provided to the evaluators were developed by the researchers of this project. No instructions were given for this form in the Carlson and Smith-Howell study.

Form C, also used in the Carlson and Smith-Howell (1995) study (see Appendix C) is a 17-item instrument which accounts for seven delivery categories (appearance, self-confidence, enthusiasm, body vitality, contact vitality, voice vitality, and speech vitality), seven content categories (evidence of thorough planning, explanations, visual aids, interest, content material, support, and logical development), and three structure categories (introduction, body, and conclusion). Each of the 17-items are rated as 0 (average), + (outstanding), or - (not satisfactory). Detailed instructions obtained from Carlson, explain what to look for when evaluating each item. The pluses and minuses are summed and the total establishes a grade as follows: +8 and above = A; +4 to +7 = B; 0 to +3 = C; -4 to -1 = D; and -5 and below = F. Half of the raters were given the written instructions that explain how to figure the grade and convert them to a percentage, while half of the raters were not given instructions.

Finally, Form D (see Appendix D), another commonly found evaluation, is currently used in an upper division undergraduate business and professional com-

munication course at a large southwestern university. This form establishes two main categories of organization and structure, and delivery. Organization and structure is comprised of seven items including: introduction, clarity of main points, support of main points, organization, transitions, conclusions, and use of persuasive elements (evidence, reasoning, emotional appeal and call to action). Delivery is established through eight items including: posture, facial expression, eye contact, gestures, composure, conversational quality, vocal delivery, and language use. Each item is measured on a five-point scale (unacceptable, poor, acceptable, good, and excellent). The general guidelines for evaluating a speech are similar to the instructions for Form B. Once again half of the raters received the instructions and half did not receive any instructions. Since no directions were available, instructions were developed by the researchers of this project.

### *Procedure*

The researchers solicited two video-taped persuasive speeches from 16 public speaking instructors. Tatum's (1992) study concluded that evaluating speeches on tape does not add to or subtract from any rater biases. The two types of speeches requested were to be "C" speech (Speech 1) and an "A" speech (Speech 2). Because the researchers attempted to minimize the influence of sex and age on the ratings, both speakers selected were female, in their early twenties and similar in appearance. Both speeches were persuasive, both were approximately the same length, and neither speaker used a visual aid. A pilot study with eight experienced raters

**Table 1**  
**Hypothesis One: Paired T-Test PRCA-24 Results for High and Moderate CAs**

<i>High CA</i>	N	Pre-test Mean	SD	Post-test Mean	SD	Change	t	p
Overall	47	88.81	6.61	74.51	11.53	14.30	8.40	.000
PS	47	26.36	2.63	20.83	4.63	5.53	9.74	.000
GRP	47	21.36	3.70	17.30	3.68	4.06	5.77	.000
MTG	47	22.53	2.85	19.70	3.61	2.83	5.17	.000
INTP	47	18.55	4.01	16.68	3.68	1.87	3.04	.004
<i>Moderate CA</i>								
Overall	268	65.38	7.97	57.38	12.77	8.00	11.57	.000
PS	268	21.02	3.66	17.89	4.69	3.13	9.97	.000
GRP	268	14.17	3.41	12.43	3.82	1.75	8.46	.000
MTG	268	15.89	3.66	14.29	3.21	1.88	7.16	.000
INTP	268	14.29	3.21	13.06	4.05	1.24	5.23	.000

was conducted in order to determine if the speeches represented the intended grade. Seven of the eight experienced raters, agreed that Speech 1 was a C-/70-72, while remaining evaluator gave the speech a D+/68. Similarly, seven of the eight experienced raters gave Speech 2 an A-/90-92, while the remaining evaluator gave the speech a B+/88.

In order to adequately assess the influence of the experience of the rater on the speech grade given, each of the 112 participants graded both speeches. The original intent was to have each group of evaluators use the same number of each form for each speech. However, some forms had to be thrown out because some evaluators erroneously received the same form twice, did not complete demographics, or failed to assign a final grade to the speech. (See Table 1.) Half of all the forms included directions on how to use the form, the other half included no directions.

Groups containing evaluators of all experience levels, met throughout a period spanning several weeks. A protocol script was followed for each group. First, the participants completed demographic forms to determine sex and age, as well as amount of speech grading experience. Next, evaluators were told they were going to see two persuasive speakers, each of whom met requirements for time limit and number of sources. Raters were told to imagine they were the speaker's instructor and the sole evaluator of the speech. Consequently, they were to evaluate the speech using the evaluation form provided for them as if they were the speaker's teacher. After the directions were given, the first evaluation form was distributed either with or without directions. Participants were given as much time as they needed to

familiarize themselves with the form. A speech was then shown (speeches were randomly ordered among groups to avoid any order effect) and raters were given as much time as they needed to complete their comments on the evaluation form. The same procedure was followed for the second speech.

Following the evaluation of both speeches, a survey was distributed to the raters. This survey consisted of open-ended questions regarding the raters' overall opinions of evaluation forms.

### ***Data Treatment***

A series of one-way ANOVAs were conducted to determine the influence of teacher experience on assigned speech grade, the influence of evaluation directions on assigned speech grade, and the affect of the evaluation form on assigned speech grade.

To get a better idea of participants' general and specific opinions about evaluation forms open-ended questions were provided on the survey. All answers were content analyzed by two trained coders who overlapped on twenty percent of the coding. The unit of analysis was a topical phrase, which was defined as a thought that can stand alone. For instance, the sentence, "I liked the form's simplicity, but I didn't like the 1-5 grading scale" was coded into two categories of "simplicity" and "grading scale problem." Scott's Pi was used to determine inter-coder reliability. The pilot coding resulted in a 95% inter-coder agreement. Final reliability was also at 95%.

## RESULTS

### *Data Analysis*

Research Question 1 investigated the influence of experience on the speech grade assigned. For Speech 1, ANOVAs revealed the mean grade of inexperienced raters was significantly higher than the grade assigned by either moderate or experienced raters ( $F[2, 112], 4.65, p = .0115$ ). Specifically, the mean for inexperienced raters was 73.86, while the means for moderate and experienced raters were nearly identical at 68.07 and 68.18 respectively. For Speech 2, the mean grade of inexperienced raters was again significantly higher than the grade assigned by either moderate or experienced raters ( $F[2, 112], 4.45, p = .0138$ ). The mean grade was 93.88 for inexperienced raters, while the means for moderate and experienced raters were 90.76 and 91.06 respectively.

Research Question 2 asked whether raters who were provided directions would grade speeches differently than raters who did not have directions. The ANOVAs indicated no significant difference between these groups and the grade assigned to either speech.

Research Question 3 explored whether evaluation forms affect speech ratings. The overall mean grades for Speeches 1 and 2 were 70.69 and 92.24 respectively. These scores fit within the projected grades for each speech. An ANOVA conducted regarding the form used and the grade given for Speech 1 showed no significant difference between the form used and the grade given. However, the ANOVA performed regarding Speech 2



revealed that Forms A, B, and C yielded similar grades, compared to Form D, which yielded a significantly lower grade ( $F[3, 112], 5.06, p = .0026$ ). The means and standard deviations were also calculated for each form and each experience level (See Table 1).

Research Question 4 asked "What are evaluator's opinions of the forms they use?" The first open-ended question asked, "Did this evaluation form include all the necessary components for you to evaluate the speech? Why or Why not? If not, what other components should be included in this evaluation form? Please explain." The answer to this question was analyzed by looking at each form individually (Forms A-D), focusing on how experienced, moderate and inexperienced evaluators answered the question. Regardless of the speech evaluated or whether or not directions were used, similar themes emerged from all three rater levels for each of the evaluation forms.

*Assessment of Form A.* In general the responses of both the experienced and moderate evaluators were positive, while the lack of directions for inexperienced evaluators yielded negative comments. Specifically, the 15 experienced evaluators who used Form A provided 22 comments in the open-ended evaluation. Only four comments were negative. Specifically, the experienced participants felt the form would be too complicated for novice evaluators, the point system was too difficult, the form was too structured, and there needed to be more specific criteria in the delivery area. The positive comments contained themes of the form being detailed (three responses), comprehensive (three responses), flexible (four responses), easy to use (four responses),

and allowing enough room to write comments (four responses).

The 16 moderate evaluators who used Form A offered 28 topical phrases. Negative comments fell into the categories of "problems with the point system" (two responses) and "form too specific" (two responses). The remaining comments were positive and included all the ideas identified by the experienced evaluators, as well as indicating that the form was easy to use (six responses) and the strengths/ weaknesses area was useful (three responses). In addition, of the participants who used directions, four indicated that the directions were useful.

The 23 inexperienced evaluators gave 26 comments regarding their use of Form A. All 10 of the negative comments were from evaluators who did not have directions to this form. These evaluators felt there was too much detail on the form (three responses), they needed more guidance with how to assign the points (five responses), and the form would be better if it had directions (two responses). The remaining 16 positive comments included the same themes as cited above and a new theme of "useful checklist" (six responses).

*Assessment of Form B.* Overall, evaluators on all experience levels had negative comments regarding Form B. Thirteen experienced evaluators, 14 moderate evaluators, and 25 inexperienced evaluators used Form B. The experienced evaluators generated 20 topical phrases about their use of Form B. The majority of these comments were negative. Experienced evaluators felt the form needed to be more specific (four responses) and to give more room to write comments (one response). Two responses were given on the need for

grading on effort or improvement. For those evaluators who had directions, concerns included: the comments not adding up to 100 percent (three responses); the poor-excellent scale seemed too arbitrary (two responses); and difficulty in matching the scale to what the evaluator thought the grade should be (one response). For those evaluators who did not have directions, three commented that directions would make the form easier to use. The remaining four comments were positive, stating that the form was flexible (two responses) and open-ended enough to "fit" any speech or speaker (two responses).

The 14 moderate evaluators also gave negative responses. Unlike their experienced colleagues, the moderate evaluators did not have a problem with the poor-excellent scale, but did cite a problem with the generality of the form. Fourteen responses claimed the scale was not specific enough. In addition, those who did not have directions stated directions would have made the form easier to use (three responses). Moderate evaluators also wanted more room to write comments (three responses). The remaining six comments were positive, citing the form's ease (three responses) and flexibility (three responses).

Finally, the 25 inexperienced evaluators echoed the others, claiming the form did not contain enough detail (12 responses) and the point system was confusing (nine responses). Again, several of those without directions stated they wanted directions for ease in grading (four responses). The six positive comments contained themes of ease of use (four responses) and flexibility (two responses).

**Assessment of Form C.** The majority of the comments for Form C were also negative for all three evaluator levels. Twenty-one experienced, 17 moderate and 23 inexperienced evaluators used Form C. Most negative comments concerned the evaluation's +,-, and 0 scoring system. Specifically, the experienced evaluators claimed the scale was confusing (five responses), complicated (six responses), and too limiting by only allowing the three options of +,-, and 0 (six responses). Other experienced evaluator comments questioned the specific grading criteria in the areas of body vitality and contact (five responses), as well as the organization of the form (four comments). Experienced evaluators also wanted more room to write comments (five responses). Evaluators not provided directions indicated that directions would have been helpful. The one positive theme was that the form was complete (four responses).

Moderate evaluators also expressed themes that the grading scale was confusing (five responses), complicated (eight responses) and limiting (four responses). They added that the grading system was too long (three responses), too subjective (three responses), and that they felt "trapped" into giving a grade they didn't want to give (four responses). On the positive side, three responses were given that the form was complete.

Inexperienced evaluators agreed with their experienced and moderate counterparts. Themes for this group included the grading scale was difficult to use (nine comments), complicated (six comments), and mathematically challenging (four comments). Inexperienced evaluators also provided the response that it would be difficult for the student to get a good grade using this form (five responses), and that the directions

were confusing (four responses). As with the other two groups, some inexperienced evaluators deemed the evaluation complete (three responses).

*Assessment of Form D.* The responses given for Form D varied among experience level. First, the 17 experienced evaluators expressed confusion with the point system (four responses), and concern that the form was not detailed enough (seven responses). They also requested more space to write comments (five responses). One participant wrote that the form "forced" him to assign a grade lower than he thought the speaker deserved. Some experienced evaluators liked Form D, saying it was flexible (three responses), complete (two responses), and gave the speaker credit for strengths rather than penalizing weaknesses (one response).

The 15 moderate evaluators agreed that the point system was confusing (nine responses), the 1-5 scale was too "constricting" (four responses), and the form was not detailed enough (five responses). There were also moderate evaluators who liked using the form, saying it was easy to use (four responses), it assisted in efficient grading (four responses), it was balanced (three responses) and it was complete (two responses).

Last, the 25 inexperienced evaluators offered a balance between negative and positive comments. They too felt the form needed to be more detailed (five responses) and have a better point system (four comments); however, they also wrote that the form was easy to use (six responses), complete (five responses), and easy for students to understand (four responses). Unlike the responses from Forms A, B, or C, none of the evaluators who used form D without directions made remarks about the absence of directions.

***Preference of Evaluation Forms.*** After the participants graded the speeches and assessed the evaluation form they used, two additional open-ended questions were posed. The first asked raters to identify, if they had a preference, which of the two evaluation forms used would be their choice to use again and why they made that choice. When given the choice between Form A and any other form, evaluators at all levels usually chose Form A. Reasons for this choice included the detailed criteria (nine responses), open-ended questions (eight responses), order of the criteria (seven responses), the checklist (five responses), space for writing (four responses), the point system (four responses) and ease of use for teachers (10 responses) and for students (four responses).

Next, when Form D was a choice with Forms B or C, evaluators regularly chose Form D. Reasons given for the choice included preference for the grading scale (eight responses), the flexibility (eight responses), and simplicity (seven responses) of the form, and the opportunity for students to get a good grade with this form (six responses — all from inexperienced evaluators). Forms B and C were seldom chosen over Forms A or D. However, when the choice was between Form B or Form C, the choice was relatively balanced for all experience levels. Form B was chosen because of its simplicity (11 responses), descriptions of each category (eight responses), and ease of use (five responses). Form C was chosen because of its thoroughness (nine responses), ease of use (six responses), and space for comments (three responses). Four experienced, two moderate, and one inexperienced evaluator did not have a preference of evaluation forms.

*Do we need evaluation forms?* The final open-ended question asked, "Do evaluators need evaluation forms in a communication course with a public speaking component? Why or why not? Please explain in as much detail as possible." The experienced evaluators offered 52 total comments, 40 of which were affirmative. Specifically, they proposed we do need evaluation forms because they help students improve (10 responses), they help students know the criteria by which they will be judged (nine responses), they help teachers remain consistent and objective among speakers (12 responses), they help justify grades (four responses), they help make grading easier (three responses), and they provide spaces for written comments (two responses). Twelve responses indicated that evaluation forms are not necessary. Eight of these responses stated simply "Teachers don't need them, but students do." The remaining four responses fell in the category of evaluation forms being not flexible enough.

Moderate evaluators gave 38 total comments, 35 of which were affirmative. They included themes that evaluation forms help with consistency and objectivity (17 responses), they give students feedback (10 responses), they assist evaluators with being organized (five responses), and offering more positive feedback (three responses). The three reasons given for not needing evaluation forms were that, "A speech is either good or bad — you don't need an evaluation form to determine that," "forms can be too stifling," and "evaluators should use their own criteria, not what is on a form."

Finally, inexperienced evaluators gave 61 comments, 56 of which were affirmative. These responses included the ideas that evaluation forms show which criteria to

grade (19 responses), and how to calculate grades (16 responses), let students know what to work on in the future (14 responses), help with consistency and objectivity (four responses), and completeness of the evaluation (three responses). Inexperienced evaluators who felt evaluation forms were not necessary stated that evaluation forms forced teachers to inflate (two responses) or deflate (two responses) students' grades.

*Types of comments given on evaluation forms.* Participants' comments on the evaluation form itself were analyzed using a coding scheme similar to the one used by Jensen and Lamoureux (1997). Specifically, the number of total comments were counted, as well as the number of positive (expresses approval) and negative (expresses disapproval or offers suggestions) comments, and the number of content and delivery comments. Attention was also paid to the evaluator's experience, which evaluation form was employed, and whether or not the evaluator was given evaluation form directions.

On average, the experienced evaluator offered 14 separate comments (topical phrases) per evaluation form; moderate evaluators offered eight comments; and inexperienced evaluators offered six comments. Form A yielded the most comments across all experience levels (11 comments), followed by Form B (eight comments), Form D (seven comments) and Form C (four comments). When moderate and inexperienced raters were provided directions, the number of comments rose on average, by four comments per evaluation. When experienced raters were provided directions, the number of comments remained consistent with the number of comments written when directions were absent. Finally, when looking at the types of comments given, experienced raters di-



rected their comments more toward content (61%) than delivery (39%), and gave more negative (59%) than positive (41%) remarks. Moderate raters offered more of a balance between content (48%) and delivery (52%), and negative (47%) and positive (53%) feedback. Inexperienced evaluators offered more delivery (68%) than content comments (32%), as well as more positive (72%) versus negative (28%) comments.

## DISCUSSION

The results of this project illustrate a variety of issues important to consider when reviewing what types of evaluation forms are used as well as who is using them. To begin, Research Question 1 (Does previous speech grading experience affect speech ratings?) was supported in the quantitative analysis. The findings show that inexperienced raters give significantly higher grades, despite the level of the speech (A or C speech). This echoes previous research that maintains training and experience are important for consistency in speech ratings (e.g., Clevenger, 1963; Bowers, 1964; Gunderson, 1978). The qualitative analysis of the evaluations written by each experience level revealed a marked difference in the amount and types of comments given. Experienced evaluators offered more comments than moderate or inexperienced evaluators, regardless of the evaluation form used. Although, as previously discussed, the number of comments fluctuated with the presence or absence of directions, the types of comments remained consistent among the three evaluator levels, regardless of the evaluation form used. Experienced

raters gave more negative, content comments; moderate raters balanced their feedback between content and delivery, and negative and positive feedback; and inexperienced raters offered more positive, delivery comments. Observing this kind of pattern clearly illustrates what issues evaluators in each experience level deems important or appropriate when grading a speech.

Research Question 2 asked if raters who have written directions on how to use evaluation forms rate speeches differently than raters who do not have directions. The data analysis revealed no differences in the speech grade given and whether or not a rater was provided with directions. However, we know from the open-ended questions that, especially when an evaluation form is complicated (like Form C, and to a lesser extent, Form A), raters like to have directions. Interestingly, as the raters' experience level increased, the requests for directions and comments about directions decreased. A qualitative analysis of the comments also revealed that, when provided with directions, moderate and inexperienced evaluators gave more written feedback to the speaker. These findings indicate that we need to continue offering our less experienced evaluators more guidance before they embark on speech grading. It appears experience enhances confidence using any evaluation form — even without directions.

The results of Research Question 3 (Do evaluation forms affect speech ratings?) show that, although a significant difference between the form used and the grade given for Speech 2 was found, no differences were found among the forms and the grade given for Speech 1. Additionally, only one form (Form D) yielded a significantly different (lower) grade on Speech 2. This indi-

cates that the evaluation form has a minimal affect on the speech rating. These findings are generally consistent with those of Carlson and Smith-Howell (1995). Further research should clarify these conflicting results.

The findings from Research Question 4 (What are evaluators' opinions of the evaluation forms they use?) reveal that, when the speech is poor, evaluators both state and demonstrate that they are more likely to follow the directions or the form. However, when the speech is of higher quality, utilizing the form as designed becomes less important to the evaluators. One possible explanation for this result is the ambiguous nature of Forms B and C. Without directions, evaluators may perceive these forms to lack any clear guidelines for grading the speech. This is further supported from the results regarding Speech 2, which indicate that Forms B and C are less likely to be followed by the rater in forming the grade. We should however note, that directions can also be ambiguous or too confining. Specifically, two experienced raters, both of whom used Form C with directions, gave extremely low grades (a 30% and a 12%) to Speaker 1. Their comments on the open-ended questions help explain these scores. One remarked "after reading the evaluation directions I felt forced into giving the speaker such a low grade." The other simply wrote "evaluation system confusing." The challenge of grading a poor quality speech, in comparison to a high quality speech was also illustrated by the much higher standard deviations on Speech 1 compared to Speech 2. This result is not surprising considering that there is only a 10-15 point range for a good speech (B to A+) compared to the possible 10-50 point range of a poor speech (B- to F). We can make two conclusions from

these results. First, when the evaluation form is complex, directions seem to be a key in using the form as it was designed. On the other hand, when an evaluation form is relatively straight-forward, evaluators can use it to determine the grade without the aid of directions. Second, it is imperative to cautiously design our evaluation forms as well as their accompanying directions so they will yield valid and reliable grades for any level of speech.

The findings from RQ4 also reveal a strong preference of certain forms over others. The answers to the open-ended questions indicate that the type of speech evaluation form used matters greatly to the rater even if the form has no bearing on the grade given to the speaker. First, examining the participants' responses shows that a balance between specificity and flexibility is wanted. That is, evaluators liked when a form offered them enough guidance to determine what was to be graded, but they didn't want to feel "forced" into giving a student a grade. Other common themes include evaluators wanting space to write comments and a logical grading system (i.e., one that added up to 100). Second, it is interesting to note the reasons cited among the three experience groups for wanting evaluation forms. The experienced raters considered them as primarily useful for students (48% of the responses), or as a way to remain consistent and objective (30% of the responses). In contrast to the student focus, moderate raters viewed evaluation forms as a useful way to stay consistent or be organized (71% of the responses). Similarly, the 70% of the comments given by inexperienced evaluators dealt specifically with the use of the form (what to evaluate and how to calculate the grade).

### ***Training Implications and Conclusion***

The finding that experienced evaluators use and think of evaluation forms differently than moderate or inexperienced evaluators is not really that surprising. However, it is noteworthy considering the number of novice TAs the basic course employs, if not to autonomously teach these courses, at least to grade speeches in these courses. Consequently, it is our responsibility as directors of these courses to provide our TAs with the tools they need to accurately grade speeches. Specifically, the current research suggests that, inexperienced raters certainly need evaluation forms which identify specific criteria to be evaluated as well as clear instructions detailing how to evaluate each item. This is particularly important for "C" speeches. These recommendations are in line with the previously cited research which established that rater training (Bowers, 1964; Gunderson, 1978), experience (Clevenger, 1963), or the combination of the two (Miller, 1964) improved the evaluation process. Considering these collective results, not only do we need well-designed evaluation forms, we also need initial and ongoing supervision as TAs learn how to most efficiently and effectively use their school's evaluation tool. For instance, considering the Jensen and Lamoureux (1997) finding that students deem specific, negative, content comments as most helpful, evaluators should be coached on *how* to use the evaluation form to determine a grade and how to write comments that will be useful to the student.

We can conclude from this study's findings that, although different evaluation forms can produce similar

grades, raters definitely have opinions and preferences regarding the form they use. Additionally, as seen in the qualitative analysis, the types and numbers of comments written to the speaker vary according to which form the evaluator employs. Because an evaluator's experience influences speech grades, as well as the amount and type of feedback given to the speaker, future research should focus on designing evaluation instruments that are more helpful to the rater as well as the student. Ensuring our own evaluation forms meet the objectives of our courses or specific assignments would be a good place for each of us to start.

## REFERENCES

- Backlund, P. (1983). Methods of assessing speaking and listening skills. In R.B. Ruben (Ed.), *Improving speaking and listening skills* (pp. 59-73). San Francisco: Jossey-Bass.
- Becker, S.L. (1962). The ratings of speeches: Scale independence. *Speech Monographs*, 29, 38-44.
- Bock, D.G. (1970). The effects of persuasibility on leniency, halo, and trait errors in the use of speech ratings scales. *Speech Teacher*, 19, 296-300.
- Bock, D.G. (1974). Reliability and validity of speech rating scales: Some error effects. *Central States Speech Journal*, 23, 145-151.
- Bock, D.G. & Bock, E.H. (1984). The effects of positional stress and receiver apprehension on leniency errors in speech evaluation: A test of the rating error paradigm. *Communication Education*, 33, 337-341.

- Bock, D.G., Butler, J.L., & Bock E.H. (1984). The impact of sex of the speaker, sex of the rater and profanity type of language trait errors in speech evaluation: A test of the rating error paradigm. *Southern Speech Communication Journal*, 49, 177-186.
- Bohn, C.A. & Bohn, E. (1985). Reliability of raters: The effects of rating errors on the speech rating process. *Communication Education*, 34, 343-351.
- Bowers, J.W. (1964). Training speech raters with film. *Speech Teacher*, 13, 228-231.
- Brooks, K. (1957). Some basic considerations in rating scale development: A descriptive bibliography. *Central States Speech Journal*, 9, 27-31.
- Carlson, R.E. & Smith-Howell, D. (1995). Classroom public speaking assessment: Reliability and validity of selected evaluation instruments. *Communication Education*, 44, 87- 97.
- Clevenger, T. (1963). Retest reliabilities of ten scales of public speaking performances. *Central States Speech Journal*, 14, 285-291.
- Clevenger, T. (1964). Influence of scale complexity on the reliability of ratings of general effectiveness in public speaking. *Speech Monographs*, 31, 153-156.
- DiSalvo, V. & Bochner, A.P. (1972). The use of simulated profiles and the Prof Technique in capturing the policies of speech raters. *Speech Teacher*, 21, 273-280.
- Gibson, J.W., Hanna, M.S., & Huddleson, B. (1985). The basic speech course at U.S. colleges and universities: IV. *Communication Education*, 34, 281-292.

- Gunderson, D.F. (1978). Video-tape modules as a device for training speech raters. *Southern Speech Communication Journal*, 43, 395-406.
- Hugenberg, L. & Yoder, D. (1996, November). A reaction to the "competent speaker speech evaluation form": An update. Paper presented at the Annual Meeting of the Speech Communication Association, San Diego. (ED 404 695)
- Jensen, K.K. & Lamoureux, E. R. (1997). Written feedback in the basic course: What instructors provide and what students deem helpful. *Basic Communication Course Annual*, 9, 37-58.
- LaLumia, J.P. (1993). Evaluating student presentations. In L.W. Hugenberg, P.L. Gray, & D.M. Trank (Eds.) *Teaching and directing the basic communications course* (pp. 241-248). Dubuque, IA: Kendall/Hunt.
- Lucas, S. (1999). Teaching public speaking. In A. Vangelisiti, J. Daly & G. Friedrich (Eds.), *Teaching communication: Theory, research and methods* (2nd ed.) (pp. 75-84). Mahwah, NJ: Lawrence Erlbaum Associates.
- Miller, G.R. (1964). Agreement and the grounds for it: Persistent problems in speech rating. *Speech Teacher*, 13, 257-261.
- Morreale, S.P., Whitney, P., Zautke, B., Ellis, K., McCormick, C. & Whitter, S. 1992, October). The center for excellence in oral communication: A comprehensive program for assessing the development of public speaking competency. Paper presented at



- the 78th annual meeting of the Speech Communication Association, Chicago, IL. (ED 350 642)
- Rubin, R.B. (1982). Communication competency assessment instrument. *Communication Education*, 31, 19-32.
- Rubin, R.B. (1999). Evaluating the product. In A. Vangelisit, J. Daly & G. Friedrich (Eds.), *Teaching communication: Theory, research and methods* (2nd ed.) (pp. 425-444). Mahwah, NJ: Lawrence Erlbaum Associates.
- Roubicek, H.L. (1990, November). Feedback: Instructional decisions to be made in performances courses. Paper presented at the Annual Meeting of the Speech Communication Association, Chicago, IL. (ED 324 722)
- Sprague, J.A. (1971). An investigation of the written critique behavior of college communication instructors. Unpublished doctoral dissertation, Purdue University, West Lafayette, IN.
- Stiggins, R.J., Backlund, P.M., & Bridgeford, N.J. (1985). Avoiding bias in the assessment of communication skills. *Communication Education*, 34, 135-141.
- Tatum, D.S. (1992, April). Controlling for judge differences in the measurement of public speaking activity. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. (ED 353 273)

## APPENDIX A

**Speech Evaluation Form A**

Name: \_\_\_\_\_ Topic: \_\_\_\_\_

Points Possible	Points Received	Check-list	Criteria
20	_____	_____	<b>Introduction:</b>
		_____	1. Captured attention
		_____	2. Stated thesis
		_____	3. Related topic to audience
		_____	4. Established credibility
		_____	5. Previewed main points
		_____	6. Provided transition to body
35	_____	_____	<b>Body:</b>
		_____	1. Organized main points clearly and logically
		_____	2. Included transitions between main points
		_____	3. Constructed effective argument for position
		_____	4. Used accurate, relevant and timely supporting materials in sufficient quantity
		_____	5. Cited sources in speech
		_____	6. Incorporated appropriate appeals to emotions, values, motivations
		_____	7. Used relevant, easy-to-see visual aids
		_____	8. Explained visual aids clearly
		_____	9. Used an oral language style appropriate to topic and audience
		_____	10. Used sound reasoning

15	_____	_____	<b>Conclusion:</b>
		_____	1. Restated thesis
		_____	2. Sumarized main points
		_____	3. Ended with a memorable final thought
30	_____	_____	<b>Delivery:</b>
		_____	1. Used adequate and inclusive eye contact
		_____	2. Used effective vocal delivery (appropriate rate and volume, clear articulation, varied inflection, and no vocal fillers)
		_____	3. Used effective physical delivery (posture, gestures, movement)

\*Major Strengths:

\*Suggested Goals for Next Speech:

\*Areas Needing Improvement:

\*Overall Evaluation:

Total Points/Grade

\*On the original form, the lower third of the page left room for these comments.

### ***Criteria for Grading Speeches-Form A***

In general, a C grade on a speech means that students have met the minimum requirements for that assignment: a grade of A or B means that students have gone beyond the minimum requirements in a significant way: and a grade of D or F means that students have failed to meet two or more of the requirements for the assignment. A grade of C represents *average, satisfactory* work. More specific information on grading criteria is provided below.

**I. A grade of C: Average, Satisfactory Work.**

**A. To be judged as average and satisfactory, the speech must:**

1. Meet all specific requirements for that speech as outlined on the assignment sheet: length, purpose, organization, research, source citation, etc.
2. Be delivered on the date assigned.
3. Address a topic appropriate to the speaker, topic, and occasion.
4. Have a full introduction and conclusion.
5. Have a clear and detectable primary purpose.
6. Include a body which has
  - a. clear and logical organization of main points.
  - b. transitions between main points.
  - c. accurate, relevant, timely and appropriate evidence and appeals in sufficient quantity.
  - d. sources of evidence cited during the presentation.
  - e. a visual aid (when necessary) which is relevant, appropriate, clearly designed and clearly explained.
7. Be delivered with adequate eye contact and animation, using a direct, conversational style.
8. Be accompanied by a sentence outline or manuscript as assigned.

**II. A grade of B: Above Average Work.**

**A. To be judged as above average, the speech must meet the criteria for a C speech, as well as:**

- 1. Exhibit skillful use of internal summaries and/or transitions.**
- 2. Demonstrate above average skill in the ability to interest and challenge the audience through the use of language, organization and supporting materials.**
- 3. Include content which shows a greater depth of research and thinking than the average student speech.**
- 4. Make a significant contribution to the knowledge or intellectual motivation of the audience.**
- 5. Involve the audience in the topic.**
- 6. Use a variety of supporting materials in an interesting and original way.**
- 7. Be delivered with poise and ease, exhibiting the personal involvement of the speaker.**

**III. A grade of A: Superior Work.**

**A. To be judged as superior, the speech must meet the criteria for a B speech, as well as:**

- 1. Constitute a genuinely individual contribution by the speaker to the thinking of the audience.**
- 2. Demonstrate exceptional skill in winning understanding of difficult concepts or**

**processes, or in winning agreement from listeners initially inclined to disagree with the speaker's ideas, or in moving an audience to action.**

- 3. Address a topic of significance.**
- 4. Include thorough research which encompasses both primary and secondary sources.**
- 5. Involved the audience throughout the entire presentation.**
- 6. Be delivered with an interesting, forceful delivery style which catches attention, motivates interest, and uses personalized directness.**

**IV. A grade of D: Below Average Work.**

- A. A speech which is below average has one or more of the following serious problems:**
  - 1. Failure to meet the basic requirements of the assignment as outlined on the assignment sheet: length, organization, research, source citation, etc.**
  - 2. Generalizations without sufficient explicit support material so that the speech material so that the speech is based only on opinion.**
  - 3. Incomplete development of ideas or lack of organization.**
  - 4. Failure to identify sources during the presentation of the speech.**

5. Reliance on only one source so that the speech is summarization of one article.
6. Superficiality which demonstrates a lack seriousness about the assignment.
7. Delivery with poor eye contact, frequent hesitations, insufficient volume, extreme dependence on notes, etc.
8. Language which evidences a written rather than an oral, style.
9. No outline.

**V. A grade of F: Unacceptable Work.**

- A. A speech which is unacceptable has one or more of the following characteristics:
  1. A majority of the problems of a below average speech.
  2. Fabricated support material.
  3. Deliberately distorted evidence.
  4. Plagiarized materials.
  5. Not presented on the assigned day.

**APPENDIX B*****Speech Evaluation Form B***

Name: \_\_\_\_\_

Topic: \_\_\_\_\_

	Poor	Fair	Average	Good	Excellent
<b>Introduction:</b> (capture attention; relate to audience; introduce topic)	—	—	—	—	—
<b>Organization:</b> (speech easy to follow; clear progression of ideas)	—	—	—	—	—
<b>Development:</b> (clear explanation; use of supporting material)	—	—	—	—	—
<b>Conclusion:</b> (provides closure; summary; vivid)	—	—	—	—	—
<b>Delivery:</b> (eye contact; understandable; use of gestures/ facial expression; conversational)	—	—	—	—	—



**Comments:**

**Rating Scale:** (A) Excellent = 90-100; (B) Good = 80-89; (C) Average = 70-79; (D) Fair = 60-69; (F) = 50-59

**Overall Rating (50-100):** \_\_\_\_\_

***Criteria for Grading Speeches — Form B***

Please rate the speaker on each category. Each category is worth 25 points. The basic criteria for each is described below.

The introduction should capture the attention of the audience, relate to the audience and introduce the topic. It should include a specific preview of main points, a thesis and a transition into the body of the speech.

The organization of the speech should be easy to follow and the progression of ideas should be clear. Although a set organizational order does not have to be followed, the organization presented should be appropriate for the topic, type of speech and audience.

The development of the topic should be clear and include supporting material. At least one source should be used per main point.

**The conclusion should provide closure. A specific review of the main points should summarize the speech. The restatement of the thesis should also be included.**

**The speaker's delivery should include eye contact, understandable vocal presentation, appropriate gestures and facial expression. The delivery should also be conversational.**

**Please do not forget to write comments for the student.**

**APPENDIX C*****Speech Evaluation Form C***

Name: \_\_\_\_\_ Topic: \_\_\_\_\_

<i>Category</i>	<i>Score (+, 0, -)</i>	<i>Comments</i>
Appearance	_____	_____
Self-confidence	_____	_____
Enthusiasm	_____	_____
Body Vitality	_____	_____
Contact Vitality	_____	_____
Vocal Vitality	_____	_____
Speech Clarity	_____	_____
Evidence of Planning	_____	_____
Explanations	_____	_____
Visual Aids	_____	_____
Interest	_____	_____
Content Material	_____	_____
Support	_____	_____
Logical Development	_____	_____
Introduction	_____	_____
Body	_____	_____
Conclusion	_____	_____
Total Score: _____ (-17 to +17 possible)		
Percentage Equivalent: _____ % Letter Grade: _____		

### ***Criteria for Grading Speeches — Form C***

Listed below are eighteen categories related to the effectiveness of a public speech. Each category is described by key words/concepts. The first seven categories relate to the speaker's delivery; the second two relate to the preparation of the speech content; the next six relate to the content as presented in the actual speech, and the final three relate to the overall speech structure.

The categories are used in grading a speech in the following manner: if the concept described by the category is average, a zero (0) is given to the category; if there is something about the elements of a category that is outstanding and significantly adds to the effectiveness of the speech, a plus (+) is given to the category; if there is something about the elements of a category that is distracting and significantly detracts from the presentation of ideas, a minus (-) is given to the category.

A philosophical assumption underlying this system is that content is most important in a speech; delivery is important only in so far as it does not detract from the content. Therefore, pluses for the seven delivery categories are extremely hard to obtain—to obtain a plus in any of the delivery categories requires that something about the delivery element significantly adds to the effectiveness of imparting the information of the speech to the audience. However, negatives for the delivery categories are relatively easy to obtain — if something about a delivery category is distracting, a minus should be given.

At the conclusion of the speech, the pluses and minuses are summed for a total score (possible scores range from -18 to +18). Grade equivalents are given as follows: -5 and less = F; -4 to -1 = D; 0 to +3 = C; +4 to +7 = B; +8 and above = A. Percentage equivalents are as follows:

-4	=	60.0%	+ 4	=	80.0%
-3	=	62.5%	+ 5	=	82.5%
-2	=	65.0%	+ 6	=	85.0%
-1	=	67.5%	+ 7	=	87.5%
0	=	70%	+ 8	=	90.0%
+1	=	72.5%	+ 9	=	92.5%
+2	=	75.0%	+10	=	95.0%
+3	=	77.5%	+11	=	97.5%
			+12	=	100%

Therefore, it is possible with this system, but extremely unlikely, to get more than 100%.

### **Delivery — Speaker Qualities**

#### *Appearance:*

- Neatness — clothing, person
- Bearing — carriage, behavior, posture
- Mannerisms — unique action or style
- Facial expression

#### *Self Confidence:*

- Composure — not agitated or disturbed
- Positiveness — definite, sure of self, forceful

#### *Enthusiasm:*

- Animation — appearance of spirit, vigor, expressiveness
- Sincerity — personally interested
- Salesmanship — punch

***Body Vitality:***

**Gestures — descriptiveness, appropriateness**  
**Purposeful movement — aimed, reasoned**

***Contact Vitality:***

**Rapport — accord, harmony**  
**Friendliness**  
**Eye contact**  
**Personality projection**

***Voice Vitality:***

**Pace, pitch, volume**  
**Projection, emphasis**

***Speech Clarity:***

**Vocabulary, grammar**  
**Articulation, pronunciation, enunciation**  
**Fillers — unmeaningful expressions**  
**Fluency — smoothness of delivery**

**Content — Preparation**

**Outline**

**Format — style, understanding, use, coordinated flow**  
**Organization — sequence, completeness, topical fit**

**Evidence of Thorough Planning:**

**Time-material relationship**  
**Continuity — smooth transitions, pointed to thesis**  
**Subject matter adequacy**  
**Audience adaptability — degree of technicality, vocabulary, etc.**

## **Content — Presentation**

### *Explanations:*

- Clarity of terms
- Completeness
- Meaningful examples

### *Visual Aids*

- Appropriateness, number, type, size
- Timeliness
- Clear explanation, handling

### *Interest:*

- Choice of subject
- Approach — humor, mood
- Interest factors — suspense, novelty, etc.

### *Content Materials:*

- Worthwhile subject — clear, concise premise
- Understanding of subject
- Adequacy of research

### *Support*

- Logical evidence
- Emotional evidence
- Use — credibility, source identification, etc.

### *Logical Development:*

- Orderly sequence — known to unknown, simple to complex
- Transitions

## **Structure**

### *Introduction*

- Gains and directs attention of audience
- Establishes speaker credibility

***Body:***

**Information relative to audience**

**Clear organization**

**Logical**

**Appropriate transitions between points**

***Conclusion***

**Summarizes major points**

**Clearly related to thesis**

**Ends with a clear, relevant statement or question**



## APPENDIX D

### *Speech Evaluation Form D*

Name: \_\_\_\_\_ Score: \_\_\_\_\_

Rating Key: 1 is Unacceptable; 2 is Poor; 3 is Acceptable;  
4 is Good; 5 is Excellent

#### **Organization and Structure:**

1. Introduction	1	2	3	4	5
2. Clarity of Main Points	1	2	3	4	5
3. Support of Main Points	1	2	3	4	5
4. Organization	1	2	3	4	5
5. Transitions	1	2	3	4	5
6. Conclusions	1	2	3	4	5
7. Use of Persuasive Elements					
Evidence	1	2	3	4	5
Reasoning	1	2	3	4	5
Emotional Appeal	1	2	3	4	5
Call to Action	1	2	3	4	5

#### **Delivery**

1. Posture	1	2	3	4	5
2. Facial Expression	1	2	3	4	5
3. Eye Contact	1	2	3	4	5
4. Gestures	1	2	3	4	5
5. Composure	1	2	3	4	5
6. Conversational Quality	1	2	3	4	5
7. Vocal Delivery					
(volume, rate, pitch, variance, etc.)	1	2	3	4	5
8. Language Use (vivid, appropriate, specificity, simplicity, etc.)	1	2	3	4	5

**General Comments:**

***Criteria for Grading Speeches — Form D***

Please rate the speaker on each sub-section. Each sub-section is worth five points, for a total of 90 points. The basic criteria for each is described below.

**Organization and Structure**

1. The introduction should capture the attention of the audience, relate to the audience and introduce the topic. It should include a specific preview of main points, a thesis, and a transition into the body of the speech.
2. The *main points* should be distinct. You should be able to easily identify them.
3. The *support* used for the main points should be complete. Evidence should be used, including, but not limited to testimony, examples and statistics.
4. The *organization* of the speech should be easy to follow and the progression of ideas should be clear. Although a set organizational order does not have to be followed, the organization presented should be appropriate for the topic and audience.
5. The *transitions* should include sentences or words to provide a bridge between the introduction and the body, between each main point, and between the body and conclusion.

6. The *conclusion* should provide closure. A specific review of the main points should summarize the speech. A restatement of the thesis should also be included.
7. The *evidence* should be cited completely and clearly during the speech. There should be a minimum of one source per main point and the information should be published within the past five years.
8. The speech should use *reasoning*. It should be logical and not contain fallacies.
9. The use of *emotional appeal* should be appropriate for the audience and the topic.
10. The *call to action* should be clearly stated steps and should illustrate a logical plan.

### Delivery

1. *Posture*: the speaker should look poised and confident.
2. *Facial Expression*: needs to be appropriate for topic and appear relaxed
3. *Eye Contact*: the speaker should frequently make eye contact all around the room
4. *Gestures*: the speaker should use gestures, but they should not be repetitive or distracting
5. *Composure*: the speaker should be confident, relaxed, polished and calm
6. *Conversational Quality*: the speaker should be well rehearsed, but not memorized or stiff
7. *Vocal Delivery*: the speaker should have appropriate volume, rate and pitch

- 8. *Language Use:* the speaker should use vivid, but appropriate imagery.**

**Please do not forget to write comments for the student.**