The Competent Speaker
Speech Evaluation Form
“The Competent Speaker”
Speech Evaluation Form and Manual, 2nd Edition

The Competent Speaker Speech Evaluation Form is a standardized and tested instrument to be used in assessing public speaking competency at the higher education level. The instrument was developed in 1990 by the NCA Committee for Assessment and Testing (now the NCA Division on Communication Assessment) and representatives of 12 academic institutions. This second edition retains the competencies and evaluative criteria from the first edition but the manual is updated to reflect more timely constructs and literature regarding the instrument and its use. The instrument can be used: (a) to evaluate informative and persuasive speeches in class; (b) for testing-in or testing-out (placement) purposes; (c) as a tool for instructing and advising students; and (d) to generate assessment data for departmental or institutional accountability.

NCA Non-Serial Publication Series
Gust Yep, Editor
San Francisco State University

The NCA Non-serial Publications (NSP) Program publishes book-length projects focusing on theoretical and/or pedagogical issues related to the study and practice of human communication in a variety of contexts. Projects grounded in social scientific, interpretive, critical, performance, and rhetorical approaches and methodologies are included. Diverse views of communication ranging from microscopic (e.g., social cognition, affect and emotion in communication) to macroscopic (e.g., public discourse, systems of representation) are also included. Topics that have been central to the history of the discipline as well as those that have been marginalized and excluded in the discipline are included as are projects with an inclusive, interdisciplinary, and social justice agenda.
# Table of Contents

I. INTRODUCTION ................................................................. 7

II. FACT SHEET FOR THE COMPETENT SPEAKER SPEECH EVALUATION FORM...................................................... 8

III. SIGNIFICANT CHARACTERISTICS OF THE SPEECH EVALUATION FORM .............................................................. 9

IV. THE COMPETENT SPEAKER SPEECH EVALUATION FORMS
   - Atomistic ........................................................................... 10
   - Holistic Form ...................................................................... 11

V. EIGHT PUBLIC SPEAKING COMPETENCIES AND CRITERIA FOR ASSESSMENT ..................................................... 12

VI. LOGISTICS, ADMINISTRATION, AND USE OF THE INSTRUMENT ..................................................................... 16
   - Training in the Use of the Form ........................................... 16
   - Using the Form to Evaluate and Grade Speeches ................. 16
   - Scoring System Options for the Form ............................... 17
   - Potential Uses of the Form ................................................ 17
   - Cautions Regarding the Use of the Form .......................... 17
   - Eight Steps for Inter-Rater(s) Reliability Training .............. 18

VII. APPENDICES, DOCUMENTATION IN SUPPORT OF THE INSTRUMENT ........................................................... 19
   - A. Background and Need for the Competent Speaker Speech Evaluation Form .............................................. 20
   - B. The Instrument Development Process ............................. 24
   - C. Reliability, Validity, and Bias Testing ............................... 26
   - D. NCA Criteria for the Assessment of Oral Communication ................................................................. 33
   - E. The Competent Speaker Compliance with NCA Criteria for Assessment .................................................. 38

VIII. BIBLIOGRAPHY .................................................................. 39
ACKNOWLEDGEMENTS

Many communication scholars and academic institutions and organizations were responsible for the original development and testing of The Competent Speaker Speech Evaluation Form. The editors of this second edition acknowledge the contributions of the following colleagues:

- The original subcommittee of public speaking experts who searched the literature and carefully developed the original speech evaluation form and its criteria. Conrad Awtry, University of Wisconsin; Jim Bradford and Faye Clark, DeKalb College; Nancy Christ, University of Nebraska; Pat Comeaux, University of North Carolina; Jean DeWitt, University of Houston; Mike Moore, Purdue University, Calumet; Sherry Morreale, University of Colorado; Phil Taylor, University of Central Florida; Karolyn Yocum, Missouri Southern State College.
- The NCA Committee on Assessment and Testing (now the NCA Division on Communication Assessment) and its chair, Phil Backlund, who organized the 1990 NCA Conference on Communication Competency Assessment, where plans for The Competent Speaker Evaluation Form were germinated.
- Statisticians Don Morley of University of Colorado at Colorado Springs and Donna Tatum of University of Chicago, who performed the initial psychometric tests of the speech evaluation form.
- University of Colorado at Colorado Springs, its communication department, faculty and staff who provided moral support and labor to help develop the instrument and its manual.
- The National Communication Association, its staff, and the NCA Non-Serial Publication Series, all of whom supported the development and testing of the original instrument as well as the production and publication of this second edition.
I. INTRODUCTION

The Competent Speaker Speech Evaluation Form was created to provide a statistically valid and reliable tool for the assessment of public speaking performance. It is an instrument designed for evaluating specific public speaking skills. Thus, the focus is on the assessment of both verbal and nonverbal behaviors characteristic of competent public speaking as opposed to knowledge about, or motivation to engage in, public speaking. The instrument can serve several purposes: (a) evaluating any type of speeches in class; (b) testing-in or testing-out (placement) purposes; (c) as a tool for instructing and advising students about the preparation and presentation of public speeches; and (d) generating assessment data for accountability-related objectives of academic institutions.

The Competent Speaker assessment instrument consists of eight public speaking competencies, four relating to preparation and four to delivery. For each of the eight competencies, a set of specific performance criteria is provided for three levels of performance: unsatisfactory, satisfactory, and excellent. Because each competency is assessed with respect to appropriateness for the audience and occasion, cultural and other biases are overcome. The eight competencies and corresponding performance criteria provide a comprehensive description of oral communication competency for the public speaking context.

Impetus for developing The Competent Speaker Evaluation Form originated during the NCA 1990 Summer Conference on Communication Competency Assessment. At that conference, a subcommittee of the Committee on Assessment and Testing was charged to:

- develop the public speaking skills portion of a test of oral communication. A position should be taken on which skills should be assessed, what criteria should be used to assess them, and suggested procedures. Perhaps prototypes should be developed and recommended. (Backlund, 1990a, p.1)

Eleven geographically-dispersed communication scholars determined this charge to be the development of a speech performance evaluation tool grounded in and driven by the communication competency and public speaking literature in the communication discipline. The subcommittee of 11 reviewed a large body of literature regarding speech evaluation and oral communication competency, synthesized that literature, and developed and tested the instrument presented in this manual. This second edition of The Competent Speaker includes the following enhancements:

- Results of new psychometric testing are presented in Appendix C.
- A holistic version of the instrument is presented immediately following the original instrument.
- The competency on supporting materials now includes electronic as well as non-electronic presentational aids.
- The academic literature and bibliography on competent public speaking are updated.

For ease of use, this manual begins with a brief fact sheet and a list of significant characteristics of The Competent Speaker. These are followed by a copy of the speech evaluation form itself, in atomistic and holistic style, and the eight competencies, including the performance criteria for each competency. Next, guidelines are provided for using the instrument to assess speakers’ performances and to train speech raters in its use. For the reader interested in the background, rationale, development process, and psychometric testing of the form, information regarding each of these issues is contained in the appendices, along with a bibliography of relevant literature.
Background

The Competent Speaker was developed in 1990 by a subcommittee of the NCA Committee on Assessment and Testing (now the NCA Division on Communication Assessment) charged to develop and test a communication competency-based speech evaluation form (Backlund, 1990). Development and testing involved representatives of 12 academic institutions, and the subcommittee was chaired by Sherwyn Morreale of the University of Colorado at Colorado Springs, Michael Moore of Purdue University, Calumet, and Phillip Taylor of the University of Central Florida. Information regarding the instrument’s development and testing in the 1990’s as well as current testing for this second edition is found in Appendices B and C in this manual.

Purpose

The instrument is to be used to assess public speaking competency at the higher education level, for purposes of in-class speech evaluation, entrance/exit placement and assessment, as an instructional strategy or advising tool, and/or to generate assessment data for institutional or departmental accountability.

The instrument assesses public speaking behaviors, as opposed to knowledge/cognition and motivation/affect. Further, it assesses molecular/specific behaviors as opposed to molar/general traits. That said, the speech evaluator may make inferences about knowledge/cognition and motivation/affect from observing the public speaking behaviors of a speaker.

Rationale

Despite the proliferation of public speaking courses and evaluation forms, no standardized and psychometrically tested speech evaluation form was available prior to 1990. Nor has there been such a form available grounded in the discipline’s conceptualization of public speaking competency. Thus, The Competent Speaker was developed to address that need for a standardized and tested speech evaluation form.

Technical Characteristics

The Competent Speaker consists of eight public speaking competencies, four of which relate to preparation and four to delivery. For each of the eight competencies, specific criteria for assessment are provided at three levels of performance: excellent, satisfactory, and unsatisfactory. The instrument is presented on pages 11 (atomistic form) and 12 (holistic form) in this manual; the criteria begin on page 13.

Conceptualization

The instrument was derived from NCA’s Speaking and Listening Competencies for High School Graduates (1982); NCA’s Communication is Life: Essential College Sophomore Speaking and Listening Competencies (Quianthy, 1990); and the public speaking competencies contained in the Communication Competency Assessment Instrument (Rubin, 1982a).

Training Manual

This manual provides guidelines for instructors to train one or more speech raters/evaluators. Training instructions for use of the instrument begin on page 21. The manual’s appendix also provides background information for administrators who may need to provide a rationale for using the instrument.

Scoring Procedure

Utilize the instrument as a ratio scale by using any numerical weighing system, between and within the eight competencies, that suits the purpose of the evaluator and the particular speech event. For example, assignment of 1 to unsatisfactory, 2 to satisfactory, and 3 to excellent, for each of the eight competencies, would result in a possible score range of 8 (unsatisfactory) to 24 (excellent) for a given speech.

Reliability/Validity/Bias

The Competent Speaker was developed with great concern for its psychometric reliability and validity and for biases of any kind and is determined to be a reliable, valid, and useful instrument with which to judge speeches. Appendix C contains the results of the original testing process as well as testing performed for inclusion in this second edition.
III. SIGNIFICANT CHARACTERISTICS OF THE SPEECH EVALUATION FORM

Every effort was made to ensure that this instrument is consistent with research and literature in the communication discipline concerning the nature of communication competence and current policy concerning its assessment. A few of the more significant characteristics of The Competent Speaker instrument relevant to this objective include the following:

1. **Assesses public speaking behavior.** In recognition of the impracticality of assessing all domains of competence, the instrument focuses exclusively on the assessment of public speaking “behaviors.” Thus, it is not designed nor does it purport to measure either knowledge/cognition about public speaking or affect/motivation for public speaking. Regardless, as with the assessment of any behavior, the instrument may permit one to make limited inferences about the knowledge, motivation, and critical thinking skills of the speaker, inferences limited at most to the particular speaking event. However, such inferences should be made with caution since any reliability or validity associated with the instrument’s assessment of behavior does not apply to its assessment of knowledge or motivation.

2. **Assesses both verbal and nonverbal behavior.** The instrument identifies competencies for verbal and nonverbal behavior and provides criteria for assessing competence in the use of each. However, assessment is limited to the public speaking context and should not be generalized to other contexts, e.g., dyadic, group, etc.

3. **Does not assess molar/general traits.** The instrument does not directly assess molar/general traits such as charisma, dynamism, audience contact, etc. Rather, the assessment of these traits is implied through the rating of molecular/specific behaviors such as communicating thesis, vocal variety, eye contact, etc.

4. **Provides a holistic assessment of remote preparation skills.** The instrument will provide a holistic assessment of the speaker’s remote preparation skills. However, additional procedures or instruments will be required to assess specific remote preparation skills (e.g., specific research or outlining skills). Of course, the results of preparation skills, as assessed in the first four competencies, can be observed to some extent as the speech is presented.

5. **Does not assess listening skills.** The instrument is designed for the assessment of “speaking” skills only. Listening skills should be assessed by instruments and procedures designed for that purpose.

6. **Provides for either an atomistic or holistic assessment.** The eight discreet competency statements with their corresponding criteria primarily provide for an atomistic/analytic assessment. However, neither the competency statements nor their corresponding criteria prevent assessment based upon a holistic impression. For this purpose, a holistic version of the instrument is included in this second edition.

7. **Assesses degree of competence.** The criteria for each competency describe degrees of competence at the unsatisfactory, satisfactory, and excellent levels. An assessment can be made of either or both the degree to which the speaker possesses each competency and the degree to which the speaker possesses public speaking competence in general.

8. **Provides descriptions of “competent” responses as anchors.** Each of the eight competency statements is accompanied by a descriptive statement of the criteria for judging each of the three degrees of competence.

9. **Is free of cultural bias.** Each competency is assessed with respect to the target audience and occasion. In other words, judgments are based upon the degree to which the behavior is appropriate to the “audience and occasion.” As long as the evaluator/assessor bases judgments on these criteria, cultural bias should not become a factor.

---

NOTE: The purchaser of this manual is permitted to copy the speech evaluation form and the criteria for the competencies presented on the following pages in order to carry out any speech performance evaluation. With the purchase of the manual from NCA, any instructor or department or institution may reprint as many copies of the speech evaluation form as are needed by that instructor, department, or institution, without further permission of NCA.
### IV. The NCA Competent Speaker Speech Evaluation Form

Course: _______  Semester: ____________  Date: ______________  Project: ___________________________

Speaker(s): ________________________________________________________
__________________________________________________________________

<table>
<thead>
<tr>
<th>PRESENTATIONAL COMPETENCIES</th>
<th>RATINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Competency One:</strong> CHOOSES AND NARROWS A TOPIC APPROPRIATELY FOR THE AUDIENCE &amp; OCCASION</td>
<td></td>
</tr>
<tr>
<td><strong>Competency Two:</strong> COMMUNICATES THE THESIS/SPECIFIC PURPOSE IN A MANNER APPROPRIATE FOR THE AUDIENCE &amp; OCCASION</td>
<td></td>
</tr>
<tr>
<td><strong>Competency Three:</strong> PROVIDES SUPPORTING MATERIAL (INCLUDING ELECTRONIC AND NON-ELECTRONIC PRESENTATIONAL AIDS) APPROPRIATE FOR THE AUDIENCE &amp; OCCASION</td>
<td></td>
</tr>
<tr>
<td><strong>Competency Four:</strong> USES AN ORGANIZATIONAL PATTERN APPROPRIATE TO THE TOPIC, AUDIENCE, OCCASION, &amp; PURPOSE</td>
<td></td>
</tr>
<tr>
<td><strong>Competency Five:</strong> USES LANGUAGE APPROPRIATE TO THE AUDIENCE &amp; OCCASION</td>
<td></td>
</tr>
<tr>
<td><strong>Competency Six:</strong> USES VOCAL VARIETY IN RATE, PITCH, &amp; INTENSITY (VOLUME) TO HEIGHTEN &amp; MAINTAIN INTEREST APPROPRIATE TO THE AUDIENCE &amp; OCCASION</td>
<td></td>
</tr>
<tr>
<td><strong>Competency Seven:</strong> USES PRONUNCIATION, GRAMMAR, &amp; ARTICULATION APPROPRIATE TO THE AUDIENCE &amp; OCCASION</td>
<td></td>
</tr>
<tr>
<td><strong>Competency Eight:</strong> USES PHYSICAL BEHAVIORS THAT SUPPORT THE VERBAL MESSAGE</td>
<td></td>
</tr>
</tbody>
</table>

General Comments:                  Summative Scores of Eight Competencies:__________
# The NCA Competent Speaker Holistic Speech Evaluation Form

## Course: _______  Semester: ___________  Date: ___________  Project: ______________________________________

## Speaker(s):

| RATINGS |
|-----------------|--------|
| Unsatisfactory | Satisfactory | Excellent |

### PREPARATION AND CONTENT
- ✔ Chooses and narrows topic appropriately
- ✔ Communicates thesis/specific purpose
- ✔ Provides appropriate supporting material (includes presentational aids)
- ✔ Uses an effective organizational pattern

### PRESENTATION AND DELIVERY
- ✔ Uses language appropriately
- ✔ Uses vocal variety in rate, pitch, and intensity
- ✔ Uses appropriate pronunciation, grammar, and articulation
- ✔ Uses physical (nonverbal) behaviors that support the verbal message

## General Comments:

## Summative Score: _______
V.  EIGHT PUBLIC SPEAKING COMPETENCIES AND CRITERIA FOR ASSESSMENT

Competency One

CHOSES AND NARROWS A TOPIC APPROPRIATELY FOR THE AUDIENCE & OCCASION.

EXCELLENT

The speaker presents a topic and a focus that are exceptionally appropriate for the purpose, time constraints, and audience.

[That is, the speaker’s choice of topic is clearly consistent with the purpose, is totally amenable to the time limitations of the speech, and reflects unusually insightful audience analysis.]

SATISFACTORY

The speaker presents a topic and a focus that are appropriate for the purpose, time constraints, and audience.

[That is, the speaker’s choice of topic is generally consistent with the purpose, is a reasonable choice for the time limitations of the speech, and reflects appropriate analysis of a majority of the audience.]

UNSATISFACTORY

The speaker presents a topic and a focus that are not appropriate for the purpose, time, the constraints or audience.

[That is, the speaker’s choice of topic is inconsistent with the purpose, the topic cannot be adequately treated in the time limitations of the speech, and there is little or no evidence of successful audience analysis.]

Competency Two

COMMUNICATES THE THESIS/SPECIFIC PURPOSE IN A MANNER APPROPRIATE FOR THE AUDIENCE & OCCASION.

EXCELLENT

The speaker communicates a thesis/specific purpose that is exceptionally clear and identifiable.

[That is, there is no question that all of the audience members should understand clearly, within the opening few sentences of the speech, precisely what the specific purpose/thesis of the speech is.]

SATISFACTORY

The speaker communicates a thesis/specific purpose that is adequately clear and identifiable.

[That is, at least a majority of the audience should understand clearly, within the opening few sentences of the speech, precisely what the specific purpose/thesis of the speech is.]

UNSATISFACTORY

The speaker does not communicate a clear and identifiable thesis/specific purpose.

[That is, a majority of the audience may have difficulty understanding, within the opening few sentences of the speech, precisely what the specific purpose/thesis of the speech is.]
**Competency Three**

PROVIDES SUPPORTING MATERIAL (INCLUDING ELECTRONIC AND NON-ELECTRONIC PRESENTATIONAL AIDS) APPROPRIATE TO THE AUDIENCE & OCCASION.

**EXCELLENT**

The speaker uses supporting material that is exceptional in quality and variety.

[That is, supporting material is unarguably linked to the thesis of the speech, and further is of such quality that it decidedly enhances the credibility of the speaker and the clarity of the topic.]

**SATISFACTORY**

The speaker uses supporting material that is appropriate in quality and variety.

[That is, supporting material is logically linked to the thesis of the speech, and is of such quality that it adds a measurable level of interest to the speech.]

**UNSATISFACTORY**

The speaker uses supporting material that is inappropriate in quality and variety.

[That is, supporting material is only vaguely related to the thesis of the speech, and variety is either too great or too little to do anything but detract from the effectiveness of the speech.]

**Competency Four**

USES AN ORGANIZATIONAL PATTERN APPROPRIATE TO THE TOPIC, AUDIENCE, OCCASION, & PURPOSE.

**EXCELLENT**

The speaker uses an exceptional introduction and conclusion and provides an exceptionally clear and logical progression within and between ideas.

[That is, the introduction clearly engages the audience in an appropriate and creative manner, the body of the speech reflects superior clarity in organization, and the conclusion clearly reflects the CONTENT of the speech and leaves the audience with an undeniable message or call to action.]

**SATISFACTORY**

The speaker uses an appropriate introduction and conclusion and provides a reasonably clear and logical progression within and between ideas.

[That is the introduction clearly engages a majority of the audience in an appropriate manner, the body of the speech reflects adequate clarity in organization, and the conclusion reflects adequately the content of the speech and leaves a majority of the audience with a clear message or call to action.]

**UNSATISFACTORY**

The speaker fails to use an introduction or conclusion and fails to provide a reasonably clear and logical progression within and among ideas.

[That is, the introduction fails to engage even a majority of the audience in an appropriate manner, the body of the speech reflects lack of clarity in organization, and the conclusion fails to reflect adequately the content of the speech and fails to leave even a majority of the audience with a clear message or call to action.]
**Competency Five**

USES LANGUAGE APPROPRIATE TO THE AUDIENCE & OCCASION.

**EXEMPLARY**

The speaker uses language that is exceptionally clear, vivid, and appropriate.

[That is, the speaker chooses language that enhances audience comprehension and enthusiasm for the speech, while adding a measure of creativity that displays exceptional sensitivity by the speaker for the nuances and poetry of meaning.]

**SATISFACTORY**

The speaker uses language that is reasonably clear, vivid, and appropriate.

[That is, the speaker chooses language that is free of inappropriate jargon, is nonsexist, is nonracist, etc.]

**UNSATISFACTORY**

The speaker uses unclear or inappropriate language.

[That is, the speaker chooses inappropriate jargon or language which is sexist, racist, etc.]

---

**Competency Six**

USES VOCAL VARIETY IN RATE, PITCH, AND INTENSITY (VOLUME) TO HEIGHTEN AND MAINTAIN INTEREST APPROPRIATE TO THE AUDIENCE & OCCASION.

**EXEMPLARY**

The speaker makes exceptional use of vocal variety in a conversational mode.

[That is, vocals are exceptionally and appropriately well-paced, easily heard by all audience members, and varied in pitch to enhance the message.]

**SATISFACTORY**

The speaker makes acceptable use of vocal variety in a conversational mode.

[That is, the speaker shows only occasional weakness in pace, volume, pitch, etc., thereby not detracting significantly from the overall quality or impact of the speech.]

**UNSATISFACTORY**

The speaker fails to use vocal variety and fails to speak in a conversational mode.

[That is, the speaker shows frequent weakness in controlling and adapting pace, volume, pitch, etc., resulting in an overall detraction from the quality or impact of the speech.]
**Competency Seven**

USES PRONUNCIATION, GRAMMAR, AND ARTICULATION APPROPRIATE TO THE AUDIENCE & OCCASION.

**EXCELLENT**
The speaker has exceptional articulation, pronunciation, and grammar.

[That is, the speaker exhibits exceptional fluency, properly formed sounds which enhance the message, and no pronunciation or grammatical errors.]

**SATISFACTORY**
The speaker has acceptable articulation, with few pronunciation or grammatical errors.

[That is, most sounds are properly formed, there are only minor vocalized disfluencies, and a few (1-2) minor errors in pronunciation and grammar.]

**UNSATISFACTORY**
The speaker fails to use acceptable articulation, pronunciation, and grammar.

[That is, nonfluencies and disfluencies interfere with the message, and frequent errors in pronunciation and grammar make it difficult for the audience to understand the message.]

---

**Competency Eight**

USES PHYSICAL BEHAVIORS THAT SUPPORT THE VERBAL MESSAGE.

**EXCELLENT**
The speaker demonstrates exceptional posture, gestures, bodily movement, facial expressions, eye contact, and use of dress.

[That is, kinesic (posture, gesture, facial expressions, eye contact) and proxemic (interpersonal distance and spatial arrangement) behaviors and dress consistently support the verbal message and thereby enhance the speaker’s credibility throughout the audience.]

**SATISFACTORY**
The speaker demonstrates acceptable posture, gestures, facial expressions, eye contact, and use of dress.

[That is, kinesic (posture, gesture, facial expressions, eye contact) and proxemic (interpersonal distance and spatial arrangement) behaviors and dress generally support the message, with minor inconsistencies that neither significantly distract from the speaker’s credibility with the audience nor interfere with the message.]

**UNSATISFACTORY**
The speaker fails to use acceptable posture, gestures, facial expressions, eye contact, and dress.

[That is, kinesic (posture, gesture, facial expressions, eye contact) and proxemic (interpersonal distance and spatial arrangement) behaviors and dress are incongruent with the verbal intent and detract from the speaker’s credibility with the audience as well as distracting the audience from the speaker’s message.]
VI. LOGISTICS AND ADMINISTRATION OF THE INSTRUMENT

Training in the Use of the Form

Before using The Competent Speaker Speech Evaluation Form to rate a public speaking performance, the evaluator first should review the entire manual, taking particular note of the Significant Characteristics of the Speech Evaluation Form on pages 9-10. Next the evaluator should thoroughly study the instrument itself and the eight competencies it contains on page 11, most importantly the criteria for each competency on pages 13-20. The evaluator also should become familiar with the descriptions of unsatisfactory, satisfactory, and excellent levels of performance for each competency.

After becoming familiar with the competencies and criteria, the evaluator should create a training disk (DVD) of six to nine exemplary speeches. Ideally, the disk should contain at least one “anchor speech” at the “unsatisfactory” level, one “anchor speech” at the “satisfactory” level and one “anchor speech” at the “excellent level.” At a minimum, the disk should contain one or more speeches in which all eight of the competencies are illustrated at the “unsatisfactory” level, one or more speeches in which all eight of the competencies are illustrated at the “satisfactory” level and one or more speeches in which all eight of the competencies are illustrated at the “excellent” level. While it is quite possible to find one speech which illustrates an “unsatisfactory” level of performance on all eight competencies, examples of “excellent” performance on each of the eight competencies may require multiple speeches. The goal in creating the training disk is to record examples of each of the eight competencies at each of the three levels of performance—unsatisfactory, satisfactory, and excellent—as a basis for establishing an appropriate level of inter-rater reliability.

The speeches on the training disk may be used to train multiple raters to a satisfactory level of inter-rater reliability by following the eight steps for such a training outlined at the end of this section on page 24. The level of inter-rater agreement may be calculated using either Cohen’s $kappa$ (Cohen, 1960), which compensates for the number of times rating categories are used, or Scott’s $pi$ (Scott, 1955), which also compensates for the rates of agreement that would be expected by chance alone.

Using the Form to Evaluate and Grade Speeches

In the actual use of the form to evaluate speeches, the evaluator may use any numerical weighing system for the three levels of competency that suits the particular context or course requirements. For example, the evaluator could: (a) examine the level of performance for each competency without assigning any numerical value to the performance of the competency and simply check off the level of performance; OR (b) assign one point for unsatisfactory, two points for satisfactory, and three points for excellent, for each competency (in this case, the range of the grade for the speech would be from 8 to 24); OR (c) for additional speeches, increase the assigned values to reflect the increasing level of importance or value of the particular speech—i.e. for each competency, assign two points for unsatisfactory, four points for satisfactory, and six points for excellent (the range for the grade would be from 15 to 48); OR (d) multiply the basic score of a speech (ranging from 8 to 24) by any number or fraction of a number to increase the total value of the speech.

In addition to using various numerical weighing systems for the three levels of competency, the evaluator also may consider differentially weighing the separate competencies, depending on the context. For example, certain competencies may be deemed more important than others in the following situations:

1. In a persuasive or research-based speech, Competency Three, “Provides Supporting Material Appropriate to the Audience and Occasion,” might be assigned more points than other competencies.
2. In a speech immediately following a lecture on style and delivery, Competency Eight, “Uses Physical Behaviors Chat Support the Verbal Message,” might be assigned more points.
3. Following a lecture on the use of language, competencies five, six, and seven might be weighed more heavily than other competencies.
4. In a persuasive speech, as opposed to an informative speech, the evaluator also may choose to weigh Competency Three higher than others, emphasizing the use of supporting material as evidence.
5. When using the form to evaluate a speech, the evaluator(s) may consider the merit and value of videotaping the speech performance. Videotaped recording and playback for students has demonstrated success in improving language usage and delivery (Dance & Zak-Dance 1986; Miles, 1981; Mulac, 1974).

**Scoring System Options for the Form**

**Option One:** Determine the point value to be awarded to each level of the eight competencies.

1. Place a simple check at the appropriate level for each competency, with no numerical value awarded.
2. Award one point for unsatisfactory, two points for satisfactory, and three for excellent for each competency.
3. As total point values for successive speeches increase, points awarded at each level for all competencies may be increased.
4. Multiply the basic speech score by any number or fraction to increase the total value of the speech.

**Option Two:** Differentially weighing the separate competencies for various speeches is also appropriate use of the form. The following suggestions may prove helpful.

1. Weigh more heavily “Competency Three: Provides Supporting Material...” for documentative or persuasive speeches.
2. Weigh more heavily “Competency Eight: Uses Physical Behaviors ...” when emphasizing style and delivery.
3. Weigh more heavily competencies five, six, and seven when emphasizing use of language.

**Potential Uses of the Form**

The Competent Speaker Speech Evaluation Form has an array of uses and applications including, but not limited to, the following: (a) entrance assessment at the beginning of a course or at the beginning of a student’s academic tenure at an institution; (b) exit assessment at the end of a course or at the conclusion of a student’s academic tenure at an institution; (c) placement assessment of a student among or into classes or courses; (d) diagnostic assessment within a class for prescriptive purposes, for providing feedback, and for encouraging development as a speaker; and (e) speech evaluation and criticism within a class.

Finally, and perhaps most important, use of this standardized speech evaluation form across sections of the same public speaking course can generate invaluable data to be used as part of assessment-related accountability programs for any department or institution.

**Cautions Regarding the Use of the Form**

Use of The Competent Speaker Evaluation Form to rate public speaking performances necessitates the observance of certain cautions. These cautions are derived primarily from the NCA Policy on Criteria for the Assessment of Oral Communication presented in the appendix to this manual and from other indicated sources:

1. The evaluator should protect the rights of the speaker/assessee by: (a) presenting the instrument, explaining how it will be administered, and defining its criteria; and (b) maintaining appropriate confidentiality of the results and their use(s) (Frey, Botan, Friedman, & Kreps, 1991).
2. The speaker’s performance and the evaluation of that performance should not be used as the “sole” instruments for procedural decisions such as placement, exemption, academic credit, or grade. Rather, they should be used only for evaluating competence in presenting a public speech. To be used for procedural decisions concerning an individual, it should be combined with other sources of information such as: (a) direct evidence of actual communication performance in school and/or other contexts; (b) results of formal competence assessment; and/or (c) measures of communication apprehension or avoidance (Taylor, 1990).
3. The evaluator using the form should be satisfactorily trained in its use as described in this manual under “Training in the Use of the Form.” For an extended discussion of evaluating speeches, refer to Evaluating Classroom Speaking (Bock & Bock, 1981).
4. The evaluator should be educated about and sensitive to the effects of relevant physical, psychological, and cultural biases (including gender, ethnic, racial, age, and developmental).
5. The evaluator should note that numerical weighing between and within competencies relates to the psychometric value of the evaluation process. Optimally, The Competent Speaker Evaluation Form is designed to be used as a ratio Scale. Such a Scale involves the assignment of numbers for the purpose of identifying ordered relations of the competencies (Williams, 1992).

6. The evaluator should avoid and be educated about rating Scale usage problems (Bock & Bock, 1981; Rubin, 1991) such as:

   (a) lack of interest, in which case the evaluator may rate the speakers inconsistently. If evaluators do not conform carefully to the criteria established for each of the competencies, they will evaluate students’ performances unfairly. Vigilance and diligence are prerequisite to fair speech evaluation. Evaluators who cannot or choose not to evaluate speeches, based carefully on the criteria of The Competent Speaker form, should be retrained or replaced;

   (b) personal bias, in which case the evaluator may be either too easy (positive leniency error) or too hard (negative leniency error) on all speakers. Or, the evaluator may be too easy (positive halo error) or too hard (negative halo error) on a specific student;

   (c) trait error, in which case the evaluator may be either too easy or too hard on a given trait (competency) on the evaluation form. This error may occur in the evaluation of one or all speakers if the evaluator attends to or neglects certain competencies; and,

   (d) central tendency error, in which case the evaluator tends to group scores toward the middle (satisfactory) range of the evaluation form. Evaluators tend to avoid making extreme evaluations, either unsatisfactory or excellent.

To avoid rating scale usage problems, adequate training and retraining in the use of The Competent Speaker Evaluation Form is encouraged. If the form is used in multiple sections of the same course, training and testing for inter-rater reliability—the degree of agreement among different evaluators or raters when judging or evaluating the same speech—is advised to provide consistency of grading between sections. Statistical tests of inter-rater reliability should be administered periodically.

**Eight Steps for Inter-Rater(s) Reliability Training**

STEP ONE: Review the entire manual, paying particular attention to the eight competencies on pages 13-20.

STEP TWO: Familiarize yourself with the assessment criteria for evaluating unsatisfactory, satisfactory, and excellent levels of performance within each of those eight competencies.

STEP THREE: View three speeches deemed unsatisfactory, satisfactory, and excellent.

STEP FOUR: Compare and discuss variations in raters’ evaluations in any of the three speeches.

STEP FIVE: Examine once again the criteria from STEP TWO as they apply to those three speeches just viewed.

STEP SIX: Evaluate the three speeches utilizing the speech evaluation form. The raters should not know the competency levels of these six speeches before evaluating them.

STEP SEVEN: Compare and discuss variations in raters’ evaluations in any of the three speeches.

STEP EIGHT: Proceed to use the speech evaluation form in the classroom or evaluation environment. Periodically, bring instructors together for a re-test of their level of inter-rater reliability.
VII. APPENDICES

The following appendix items clarify the process by which the original Competent Speaker Speech Evaluation Form was researched, developed, and tested from 1990 to 1993 as well as more recent literature and testing for this second edition of the instrument and manual. These appendix items are supplied to the rater or administrative user to support and inform the selection and use of this instrument. If any further supportive information is required by an accreditation agency or academic institution for validation or administrative purposes, that support is available through any of the authors of this manual:

A. Background and Need for the Competent Speaker Speech Evaluation Form
B. The Instrument Development Process
C. Reliability, Validity, and Bias Testing
D. NCA Criteria for the Assessment of Oral Communication:
E. The Competent Speaker Compliance with NCA Criteria for Assessment
APPENDIX A

BACKGROUND AND NEED FOR THE COMPETENT SPEAKER SPEECH EVALUATION FORM

Historical Background

An examination of academic literature, as far back as the 1970s and through the first part of the 21st century, reveals the degree to which “communication competence” has become a significant referent with respect to the goals of communication instruction. While earlier efforts focused on defining and describing what it means to be a competent communicator, later work tended to apply competence constructs as variables in empirical studies.


The availability of new technology since the initial publication of The Competent Speaker in 1993 opened up new avenues for empirical work and measuring speech competence but with less focus on clarifying components of competence (Cronin, et al, 1994; Hinton & Karmer, 1998; Sawyer & Behnke, 2001). In articles published in a special 2002 edition of Communication Education, various authors noted that the variety of approaches and content in communication curricula actually pose a challenge for clarifying and narrowing of competencies (Morreale & Backlund).

Further examination of these and other sources suggests that communication competence and assessment are intrinsically intertwined. Partly as a response to accreditation and legislative mandates, educators and administrators perceive a need to clarify and describe the nature of communication competence but, equally important, to assess students’ achievement of it. In many, if not most cases, the communication competence of greatest interest continues to be public speaking (Morreale, Hugenberg, & Worley, 2007).

The Need for Assessing Communication and a Standardized Public Speaking Evaluation Instrument

The need for a standardized public speaking assessment instrument has become increasingly evident as research concerning communication competence generally and the assessment of communication competence specifically has progressed in recent years.

The strongest statement by the National Communication Association (NCA) and communication educators concerning the importance of communication competence and the need for appropriate assessment strategies was provided both by the convening and the results of the 1990 Summer Conference on the Assessment of Oral Communication Skills. Among the many issues surrounding communication competence and its assessment addressed by the conference participants, one issue concerned the development of a “public speaking skills portion of a test of oral communication” (Backlund, 1990a, p. 5). This was the charge given to one group of conference participants. This group, consisting of communication scholars from 11 universities throughout the United States, was further charged to take a position “on which skills should be assessed, what criteria should be used to assess them, and suggested procedures” (Backlund, 1990a, p. 5). The scholars were asked to consider developing and recommending appropriate assessment prototypes. The efforts of this group resulted in the development, testing, and publication of the first edition of The Competent Speaker by NCA in 1993 and this subsequent second edition.

This review now discusses three issues relevant to the conceptual foundation in communication competence of The Competent Speaker, 2nd Edition. Each of these three issues was considered by the developers of this instrument as they determined what aspects, skills, or public speaking behaviors to assess.
Relevant Issues

Despite the many issues of controversy surrounding the nature of communication competence and its assessment, there are three central issues that have emerged about which there is increasing agreement.

1. The first issue concerns the nature of the components of communication competence. The literature reveals a great degree of controversy over the years concerning this question. The primary controversy concerns which of three broad components, corresponding to Bloom’s (1964a, 1964b) taxonomy of cognitive, affective, and psychomotor domains, comprise communication competence. Of these, there appears to be consistent agreement that competence consists of at least knowledge or cognition and behavior (Backlund & Wiemann, 1978; Cooley & Roach, 1984; Fogel, 1979; Harris, 1979; McCroskey, 1982, 1984; Spitzberg, 1983; Wiemann & Backlund, 1980). In other words, communication competence requires both a repertoire of skills and a body of knowledge concerning those skills and their implementation.

The affective domain has increasingly been included as a third component. The position here is that communication competence consists not only of knowing what and how to perform, but also valuing performance sufficiently to do so (Moore, 1981). This component, more frequently referred to as motivation, was identified by Rubin (1983 as the third dimension of competence, by Spitzberg and Hecht (1984) as integral to their model of relational competence, and by Spitzberg and Cupach (1989) as a component of interpersonal competence. Moreover, to the degree that such predispositions as communication apprehension (McCroskey, 1970, 1977), receiver apprehension (Wheeless, 1975) and willingness to communicate (McCroskey & Richmond, 1987; Mottet, Martin & Myers, 2004) are related to communication competence, the role of motivation, or the affective domain, is well supported.

While there appears to be general consensus that communication competence consists of knowledge, skills, and motivation, there is much less clarity concerning the specific sub-components which comprise each of these domains. Typically, researchers have attempted to identify the specific communication behaviors or skills which comprise competence. This has resulted in an extensive and highly diverse list of behaviors, from broad macro-behaviors such as empathy, behavioral flexibility, interaction management, listening, and speaking, (Bochner & Kelly, 1974; DiSalvo, 1980; MacIntyre & MacDonald, 1998; Step, M. M. & Finucane, M. O., 2002; Wiemann, 1977) to more specific micro-behaviors such as articulation and pronunciation (Duran, 1983; Rubin, 1982). In the 1998 NCA document Speaking and Listening Competencies for College Students (Morreale, S. P., Rubin, R. B., & Jones, E. A., 1998), two sets of generalized skills describe communicative behavior and competencies expected for the college undergraduate. A summary of those expectations for undergraduates was later published by the same authors (Rubin & Morreale, 2000), and an encyclopedia on communication includes a similar description of the nature of student communication competence (Morreale, 2007).

These recent efforts aside, Spitzberg (1987) points out that, “Given a lack of conceptual guidance, measurement efforts to date have yet to identify a comprehensive or consistent set of competence components” (p. 5). This relates directly to the second issue central to the nature of communication competence.

2. Is communication competence a trait or a state? At one level, the question here is whether communication competence is a cross-situational disposition or whether it is dependent upon the situation. Trait research expanded in the last ten years from work in psychobiology but is still limited primarily to interpersonal trait behaviors (Beatty, Heisel, Hall, Levine, & LaFrance, 2002; Beatty, Marshall, & Rudd, 2001; Beatty & McCroskey, 1997, 1998; Beatty, McCroskey, & Heisel, 1998; Beatty, McCroskey, & Valentir; McCroskey, 2001; Heisel, LaFrance, & Beatty, 2003; McCroskey, Heisel, & Richmond, 2001; Valentir, Beatty, Rudd, Dobos, & Heisel, 1998; Wahba & McCroskey, 2005). Research in public speaking anxiety has expanded in other fields including virtual reality (Pertaub, Slater, Steed, 1999, social anxiety behavioral research (Harb, Eng, Zaider, Heimberg, 2003), and dealing with the fear of public speaking through Internet-based therapy (Botella, C., Banos, R., Guillen, V., et al., 2000).

While the research suggests a few traits, such as rhetorical sensitivity and communication apprehension, exhibit cross-situational consistency, most research indicates that competence appears to be too situationally bound, and that the research is too fraught with methodological problems to posit communication competence
The Competent Speaker Speech Evaluation Form

as a general disposition (Andersen, 1987; Rubin, 1990; Spitzberg, 1987). This is further supported by the quantity of research concerned with identifying competent communication behaviors in such diverse settings as interpersonal (Andersen, 1987; Bochner & Kelly, 1974; Parks, 1985; Spitzberg & Hecht, 1984), group (Bradley, 1980), public (Quianty, 1990), organization (DiSalvo, 1980; Monge, Bachman, Dillard, & Eisenberg, 1982; Papa, 1989), mass media (Anderson, 1983), gender (Sellnow & Treinin, 2004) and intercultural (Chen, 1988; Cooley & Roach, 1984). As Rubin (1990) states, “Research continues to point to a need to use both state and trait measures to examine communication competence until we have a firm understanding of which measures assess traits and which estimate state-influenced behaviors” (p. 104). This is consistent with Spitzberg and Cupach’s (1989) view that “both trait and state approaches to the conceptualization of competence seem viable and even compatible” (p. 53).

3. The third issue central to this discussion is frequently characterized as the “effectiveness vs. appropriateness” debate. While there appears to have been general consensus that appropriateness is directly related to competence, the relationship between competence and effectiveness has not been as consistently clear. Some conceptualizations are not very explicit about the relationship between effectiveness and competence. Others, however, state quite explicitly that effectiveness is a fundamental criterion of competence.

Representative of those who are not explicit about the relationship between effectiveness and competence are Allen and Brown (1976), who view communication competence as “an awareness of the transactions that occur between people” (p. 248). While this perspective ties competence “to actual performance of language in social contexts” (p. 248), nowhere in the explication of the four principal features of competence is it clearly tied to effectiveness. Similarly, Harris (1979) is less specific about the role of effectiveness in her definition of interpersonal competence as “the ability to create and coordinate interpersonal systems” (p. 32). McCroskey (1982), however, is quite explicit in his statement that communication effectiveness “is neither a necessary nor sufficient condition for a judgment of competence” (p.3). Bochner and Kelly (1974) and Heath (1977) suggest only a general link between effectiveness and competence, but others are much more specific. Weimann (1977), for example, defines communication competence as: the ability of an interactant to accomplish goals “while maintaining the face and line of his fellow interactants within the constraints of the situation” (p. 198). Fogel (1979) expresses an even stronger link between effectiveness and communication competence by defining the latter as the ability “to affect another’s attitudes, behaviors, and beliefs” (p. 15). Spitzberg (1983), on the other hand, argues that both appropriateness and effectiveness are essential to competent communication.

Despite the different emphases of various definitions, both effectiveness and appropriateness appear to have a role in communication competence. As Rubin (1990) explains, “Early distinctions between competence (appropriateness) and effectiveness (goal accomplishment) seem to have faded. Most researchers today agree that both elements must be present” (p. 109).

There appears to be widespread agreement that communication competence consists of at least three dimensions (cognitions/knowledge, affect/motivation, behaviors/skills), is both a general disposition and context dependent, and requires behavior which is both effective and appropriate. Each of these issues has implications for any assessment of communication competence.

Conclusions and Ramifications for The Competent Speaker Form

To summarize, communication is competent, which means of high quality, when it is both appropriate and effective for the particular situation (Morreale, Spitzberg, & Barge, 2006). Appropriate communication means that you act in ways suitable to the norms and expectations of the context and situation in which you find yourself. Effective communication means you are able to achieve the most desirable objectives or outcomes in the context.

In order to communicate competently, there are three basic requirements you must meet. First, you must be motivated to communicate competently. Second, you must be knowledgeable about the situation in which you are communicating and the kind of communication expected and needed in that situation. Third, you must be skilled at actually sending and receiving messages in that particular situation. These three requirements or dimensions of competence – motivation, knowledge, and skills – are the foundation of competent communication whether you are
in an interpersonal situation, a group, public speaking, or even in a mass communication context such as on television or in a mediated context like using e-mail.

Perhaps the clearest and most important implication of these conclusions is the virtual impossibility of developing a single instrument which can provide a valid and reliable assessment of communication competence. The first barrier is provided by the multi-dimensional nature of competence. While the cognitive and affective dimensions potentially could be assessed by a single paper-and-pencil instrument, the behavioral dimension could not — requiring a two-part instrument/process at a minimum. However, the second barrier, the contextual nature of communication competence, is insurmountable. Again, while it may be possible to develop a single instrument to assess the cognitive and affective dimensions of the primary communication contexts (dyadic, group, public, mediated, etc.), a single instrument or procedure to assess the behavioral dimension of all contexts would not be possible.

Given the impracticality of developing a single instrument to assess communication competence, the focus must be on developing multiple instruments or procedures for assessing competence within specific contexts. One of the most salient contexts for most speech communication educators and classrooms, at all levels, is the public speaking context.

While there are numerous instruments available for evaluating public speaking performance, Most do not have a clearly established basis in theory, nor have they been systematically tested for validity and reliability with the exception of Rubin’s Communication Competency Assessment Instrument (1982). However, this instrument is designed to provide a more comprehensive assessment of communication competence, of which public speaking is one part, and is designed for use outside of the typical classroom setting.

Based on the dearth of a standardized and tested public speaking evaluation form, The Competent Speaker Speech Evaluation Form, grounded in the communication competency paradigm, was developed. Of necessity, the developers narrowed their focus to assessing the skills dimension of one context – public speaking. As reported in the development section of this manual, their efforts focused on determining which particular skills most adequately support perceptions of public speaking as appropriate and effective.
APPENDIX B

THE INSTRUMENT DEVELOPMENT PROCESS

For this, the second edition of The Competent Speaker Evaluation Form, only one minor modification was made by the instrument developers to the substance of the eight competencies. In response to requests from users of the instrument, the developers made one addition to Competency Three, which focuses on the use of supporting materials. Over the years, users were unclear as to which competency should subsume the use of electronic presentational aids such as PowerPoint. The developers added the phrase “including electronic and non-electronic presentational aids” to Competency Three.

The following narrative outlines the process by which the eight original competencies and their descriptors were developed.

Historical Development of the Competent Speaker Form

After considerable deliberation, the subcommittee of the Committee on Assessment and Testing decided that the greatest need was for an instrument that could be used for the following purposes: (a) as a pedagogical tool for the evaluation and development of public speaking skills in the classroom; and (b) as a pre (testing out) and/or post (exit) assessment of public speaking skills in the basic and/or public speaking course. This further suggested that the instrument should be developed for the college sophomore (grade 14) level. This would provide an instrument that could be used as a post (exit) assessment at the high school (grade 12) level, as well as either a pre (testing out) or a post assessment at the college sophomore level.

The subcommittee also decided that the instrument should be based upon, and an extension of, the research concerning both communication competence and public speaking assessment. After an extensive review of the literature, outlined earlier in Appendix A and in this manual’s bibliography (Moore & Awtry, 1991), the two documents which appeared most germane to the content of a public speaking assessment instrument were NCA’s Speaking and Listening Competencies for High School Graduates (1982 and NCA’s Wingspread Conference Proceedings, Communication: is Life: Essential College Sophomore Speaking and Listening Competencies (Quianthy, 1990). The task force selected the public speaking competencies identified in the “college sophomore” document (Table 1) and reviewed them against the “high school” document (Table 2) to insure that the competencies subsumed all of the public speaking competencies listed for the high school graduate. Satisfied that the list of competencies represented the best and most current thinking within the discipline concerning the characteristics of public speaking competence, the subcommittee agreed to modification in the rewording of the competencies to insure clarity of language and consistency of structure. This resulted in the eight competency statements which comprise The Competent Speaker instrument (Table 3).

Table 1

<table>
<thead>
<tr>
<th>Speaking Competencies Listed in Communication is Life: Essential College Sophomore Speaking and Listening Competencies, (Quianthy, 1990).</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Determine the purpose of oral discourse.</td>
</tr>
<tr>
<td>- Choose a topic and restrict it according to the purpose and audience.</td>
</tr>
<tr>
<td>- Fulfill the purpose of oral discourse by formulating a thesis statement, providing adequate support material, selecting a suitable organization pattern, demonstrating careful choice of words, providing effective transitions, and demonstrating suitable interpersonal skills.</td>
</tr>
<tr>
<td>- Employ vocal variety in rate, pitch, and intensity.</td>
</tr>
<tr>
<td>- Articulate clearly.</td>
</tr>
<tr>
<td>- Employ the level of American English appropriate to the designated audience.</td>
</tr>
<tr>
<td>- Demonstrate nonverbal behavior that supports the verbal message.</td>
</tr>
</tbody>
</table>

24 The Competent Speaker Speech Evaluation Form
Table 2

Speaking Competencies listed in NCA Guidelines:
Speaking and Listening Competencies for High School Graduates, (1982).

- Use words, pronunciation, and grammar appropriate for situation.
- Use nonverbal signs appropriate for situation.
- Use voice effectively.
- Express ideas clearly and concisely.
- Express and defend with evidence your point of view.
- Organize (order) messages so that others can understand them.
- Summarize messages.

Table 3

Eight Public Speaking Competencies listed in The Competent Speaker,
(Morreale et al., 1990).

- Chooses and narrows a topic appropriate to the audience and occasion.
- Communicates the thesis/specific purpose in a manner appropriate for audience and occasion.
- Provides supporting material appropriate to the audience and occasion.
- Uses an organizational pattern appropriate to the topic, audience, occasion, and purpose.
- Uses language appropriate to the audience and occasion.
- Uses vocal variety in rate, pitch, and intensity, to heighten and maintain interest appropriate to the audience and occasion.
- Uses pronunciation, grammar, and articulation appropriate to the audience and occasion.
- Uses physical behaviors that support the verbal message.

Following identification of the eight public speaking competencies, the task force developed performance standards (criteria) by which each competency could be evaluated or measured (see The Competent Speaker: Eight Public Speaking Competencies and Criteria for Assessment beginning on page 10 of this manual). As with the competencies, the criteria were based upon the competency literature identified earlier, as well as upon published guidelines for speech evaluation (see, for example, Bock & Bock, 1981; Mead & Rubin, 1985; Powers, 1984).

As explained by Morreale (1990) in an earlier report, following development of the eight competencies and corresponding evaluative criteria, the task force generated a speech performance evaluation form, The Competent Speaker Speech Evaluation Form (see the instrument itself on page nine of this manual). In designing the evaluation form, earlier published NCA guidelines for constructing a speech evaluation instrument (Bock & Bock, 1981, 21-22), were considered, with particular emphasis on Scale construction and controlling for rater errors. Additionally, National College Board recommendations for the development of measures of speaking, and listening (Powers, 1984, 7), were observed. Moreover, every effort was made to insure the instrument conformed to the “NCA Policy on Criteria for the Assessment of Oral Communication,” a policy statement generated at the 1990 NCA Summer Conference on the Assessment of Oral Communication (Cracker-Lakness, 1991). (See Appendix D for the NCA policy statement and Appendix E for a notated list of those criteria with which The Competent Speaker instrument complies). Two of the more important criteria, those concerning the instrument’s reliability and validity, are addressed in Appendix C to follow.
APPENDIX C

RELIABILITY, VALIDITY, AND BIAS TESTING

Traditional Analysis

The following section describes the testing processes from the original study using traditional statistical analyses to determine the reliability and validity of The Competent Speaker Evaluation Form. Further investigation for this second edition then is reported in the Rasch Analysis section that follows.

Development of a Testing Videotape

In order to test the reliability and validity of The Competent Speaker Evaluation Form and criteria, a videotape was developed with student speeches in an actual classroom environment at a Midwestern University. The student speeches were informative presentations lasting approximately five minutes in duration. A group of six Graduate Teaching Assistants (GTAs) selected approximately 40 videotaped classroom speeches from the previous semester’s presentations. Utilizing the form and criteria, the GTAs rated these speeches as either unsatisfactory, satisfactory, or excellent. From this initial pool of 40, the public speaking course director and one of the GTAs selected 12 student speeches, four at each level of competency. These 12 speeches were transferred to one master tape for training and rating purposes. The student sample represented on the rating video was mixed by gender (five females and seven males) and by ethnicity (nine Whites, one Black, one Hispanic, and one Filipino). The 12 speeches then were randomly ordered 12 different times, utilizing a table of random numbers, and placed on tapes that were sent to 12 speech communication professionals at 12 different universities. Additionally, one speech exemplary of each level of competency was selected as an anchor by which raters could become familiar with performance at each level of competency before rating the sample of 12 speeches. The three anchor speeches were placed at the beginning of the master tape and were identified as to the level of competency each represented.

Raters

The 12 raters were speech communication professionals teaching at colleges and universities in the U.S. Nine of the raters held a Ph.D. or equivalent, while three of the raters held master’s degrees. The raters’ experience in teaching ranged from 4 years to 25 years as estimated by the date of receipt of the raters’ terminal degrees. Raters were eight females and four males, eleven of which were Anglo and one Hispanic.

Raters received a packet containing instructions for self-training on the use of the speech evaluation form and criteria and the tape with the 12 student presentations. Specifically, the raters were instructed to: (a) review the standards and criteria for the competencies before viewing any speeches; (b) view the three exemplary speeches while simultaneously reviewing the standards and criteria; and (c) view each of the 12 speeches without making any formal evaluation, review the standards and criteria as they pertained to that speech, and finally, view the speech one more time and enter the evaluation on the rating form.

Overall inter-rater reliability for the students’ total score on the instrument was high for the 12 raters with Ebel’s (1951) coefficient reading .92. Inter-rater reliability was also examined for each of the eight competencies. The 12 raters achieved a high degree of reliability on the eight competencies with Ebel’s coefficient ranging from .90 to .94.

In addition to using 12 speech communication professionals to test reliability, other reliability testing was conducted utilizing 10 GTAs as raters. The raters were from two Midwestern universities. Half of the GTAs had utilized the speech evaluation form for one semester and the other half were given a brief training in the instrument’s use. An inter-rater reliability test for the GTAs generated a Cronbach coefficient of .76.

In addition to the GTAs, inter-rater reliability testing was conducted with a small pool of community college speech instructors (N=3). They received a brief training with the instrument before evaluating the 12 videotaped speeches. This inter-rater reliability test generated a Cronbach coefficient of .84.
Validity

In addition to reliability testing of The Competent Speaker form, the instrument and its criteria currently meet content or face validity. That validity can be argued based on the extensive literature review conducted during the process of development of the instrument by the 11 member subcommittee of the NCA Committee on Assessment and Testing (see Appendices A and B of this manual).

Also regarding the validity of the instrument, two studies testing for convergent validity have been conducted. First, a correlation of scores on the public speaking portion of the Personal Report of Communication Apprehension (McCroskey, 1970) and scores derived using The Competent Speaker form in an introductory speech class indicated inverse directional convergent validity. That is, as speech scores using The Competent Speaker form increased, scores on the public speaking items of the PRCA decreased. Second, a correlation of scores on the seven public speaking items of the Communication Competency Assessment Instrument (Rubin, 1982), derived from entrance and exit interviews with students in an introductory speech class, were correlated with The Competent Speaker scores from the same class. Positive directional convergent validity for the two instruments was indicated; scores on speeches, rated using The Competent Speaker and scores or, the public speaking items of the CCAI both increased.

Cultural Diversity

In addition to The Competent Speaker Evaluation Form and criteria undergoing reliability and validity scrutiny, several other tests evaluated the form in regard to ethnic and gender bias. One study compared the 12 speech communication professionals’ ratings of 12 speeches to the ratings of the same speeches by a sample of 28 minority students using the speech evaluation form. An inter-rater reliability test of the minority students as a group generated a Cronbach’s alpha coefficient of .76. When combined with the 12 speech communication professionals, the coefficient remained .76. In another two-pronged diversity study, actual speech evaluations in the classroom (N=260) were examined both by ethnicity and by gender. In an analysis of variance, no significant difference was found in the ratings of speeches of White (m=86.46), Black (m=82.92), Hispanic (m=85.25), or Asian (82.33) students (F=.16). Additionally, there was no significant difference in speech scores of female (m=86.47) or male (m=85.70) students using the speech evaluation form (F=.41).

Normative Data

Although normative data are provided for training purposes with the videotaped anchor speeches, The Competent Speaker Evaluation Form and standards are criterion referenced. The competencies and criteria were developed based upon the literature investigated.

Conclusions of Traditional Analysis

Initial and subsequent testing of The Competent Speaker Evaluation Form indicates that the instrument is psychometrically sound in terms of reliability and validity.

As described in the following section, larger and more diverse pools of raters and greater diversity of student populations are evaluated for this edition of The Competent Speaker Evaluation Form. These studies are found to be confirmatory of the original results.

Rasch Analysis

After researchers develop an instrument expected to measure a phenomenon, they test it in the field. Data analysis of the collected sample(s) determines whether or not the instrument is deemed a valid mode of measurement.

In 1953 Georg Rasch, a Danish mathematician, developed a unique model for item analysis. This method produces results that are distinctly different from traditional statistical analysis. A statistical analysis describes a one-time event. The elements of the event are inextricably bound together into one observation. Those elements are, in this case, the items on the evaluation form, the raters using it, and the speeches they are judging.

The results of a traditional statistical analysis are not generalizable or comparable across samples or time. However, the unique feature of the Rasch model is that it allows the researcher to separate the elements under investiga-
tion and focus on one at a time. They are independent of each other and are all measured in common units, “logits.” Instead of using the “rubber ruler” of statistics, we can now use the Rasch model to condition those raw scores into meaningful measures that are calibrated upon a line of inquiry.

**Data Analysis**

The computer program FACETS 3.56 provides the basis for this analysis. It uses an extension of Rasch’s original separability theorem. John Michael Linacre, formerly of the MESA Psychometric Laboratory at the University of Chicago, and currently at the University of the Sunshine Coast in Australia, generated the model for many-faceted conjoint measurement. Once raw scores are conditioned into measures, traditional statistical analyses are performed with SPSS 13.

The Rasch Model for Conjoint Measurement:

\[
\frac{P_{njik}}{P_{njik} - 1} = B_n - C_j - D_i - F_k
\]

- \(B_n\) \(n = 1\) to \(n\) (speakers/speeches)
- \(C_j\) \(j = 1\) to \(n\) (judges)
- \(D_i\) \(i = 1\) to \(n\) (items)
- \(F_k\) \(k = 1\) to \(n\) (rating scale categories)

This method allows one to examine the various elements in an assessment situation. All of the facets are calibrated in common units of measure within a common frame of reference. An objective measurement analysis performs the following functions:

1) provides a calibration of evaluation items (difficulty)
2) produces objective measures of speakers’ competence
3) measures the severity of the raters
4) discovers rater inconsistency
5) detects rating scale step structure

**Data Description**

The original study for the first edition of *The Competent Speaker* began with 12 speech teachers rating the 12 speeches on the Testing Video referenced above. A nine-point rating Scale was used to judge the level of ability for each competency. This comprised the data for the initial analysis. In subsequent analyses, graduate assistants, community college teachers and undergraduate students were added and judged all 12 speeches using a three-point rating Scale.

The current study has much more data to investigate. Since *The Competent Speaker Evaluation Form* was introduced, many schools have successfully used it. Four universities and a community college shared samples of their speech ratings. Some speeches were rated on videotape; most were rated in class. Teachers, classmates and independent raters used the form to assess speeches. Some forms used a six-point rating scale; most used a three-point scale. Raters and speakers reflect the diversity found across the country.

Analyses are performed on individual datasets as well as various combinations of data. The complete database used in this study consists of:

- Number of Speeches: 583
- Number of Speakers: 294
- Number of Raters: 328
- Number of Rating Forms: 8,945
**Rasch Measurement Reports**

When reading the Rasch measurement reports, all numbers are directly comparable. For example, money is in common units; we all know there are 100 pennies in a dollar and that a “dollar” is a “dollar.” The same frame of reference applies to investigating and measuring public speaking competence. When measurement reports are given, they are in units of measure called “logits” (for log odds unit). Each logit has 100 points and has the same properties as a dollar. We can compare one “logit/price” to another. We can add and subtract with logits which are stable, standardized units of measure. Person A’s first speech measure is 10.05, and her third measure is 11.45. We know she has improved her public speaking by 1.40 logits, or 140 points.

The scales have been calibrated so the origin, or balance point is “10.00.” This means an evaluation item that is of average difficulty; or a person who is of average ability; or a rater who is of average severity has a measure of 10.00. The lower the number, the easier the item; or the person is less able, or the rater is less harsh. When a measure is higher than 10.00 it indicates a harder item; or more ability or more severity than that of the “average” person.

A metric is established and maintained that can be used from one time to another, or situation to situation. This provides the method to establish a benchmark to track and assess improvement in public speaking.

**The Evaluation Form**

The first step in a Rasch analysis is to determine if the evaluation form defines the variable “Public Speaking Competence.” Is it a valid instrument for raters to use when judging speeches?

All items are not created equal. That is, a range of difficulty must be covered if a test is to be useful in measuring any variable. The “ruler” that measures the variable must be calibrated in equal units, or “logits.” The items that comprise this ruler lie upon a line of inquiry that is centered on “ten”. Items with calibrations below 10.00 are easier to accomplish; those with calibrations above 10.00 are harder to achieve.

If the hierarchical arrangement of the items along the line of inquiry defines the variable and corresponds with the intention of the study, then construct validity is established. Examining each item’s content in relation to its calibration, fit statistics, and point biserial correlation reveals content validity.

**The Competent Speaker Speech Evaluation Form** is demonstrated as a valid, useful instrument with which to judge speeches. The following table, based upon an analysis combining all of the data, shows items defining the variable, “Public Speaking Competence.” The items are listed in order of difficulty, from easiest to hardest.

<table>
<thead>
<tr>
<th>Measure</th>
<th>is the item’s calibration – its placement upon the line of inquiry – expressed in logits. Error is how accurately the item is measured. The items cover a reasonable range, which shows the form does a good job of defining public speaking competency.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean square (MNSQ)</td>
<td>represents the expected response to the item and its fit along the line of inquiry. If any item has a high mean square, then there is a question of whether it fits upon the line of inquiry and is helpful in defining the variable. A mean square of 1.0 is expected; .7 to 1.3 is normal. If the mean square is outside the range, the item is behaving erratically and needs to be examined more closely. For example, if an item has a mean square of 1.6, that means there is 60% more noise than what is expected; 2.4 is 140% more than expected. The standardized fit statistic (similar in nature to a t-test) is the sureness of the item’s fit.</td>
</tr>
<tr>
<td>Reliability reflects the degree to which scores are free of measurement error. A test, or evaluation form that produces highly consistent, stable results (e.g., relatively free from random error) is said to be highly reliable. The higher the value of the index (closer to 1.0), the greater is the reliability. Every data analysis, whether it is an individual class or combined datasets, has a Cronbach alpha of .90 or higher. These results give great confidence that using The Competent Speaker Speech Evaluation Form produces reliable assessments.</td>
<td></td>
</tr>
</tbody>
</table>
The items on the **Competent Speaker Evaluation Form** fit the pattern of expected responses. This means all the items are on the line of inquiry and contribute to the definition of the variable.

### Table 1

<table>
<thead>
<tr>
<th>Item</th>
<th>Measure</th>
<th>Error</th>
<th>MNSQ</th>
<th>PtBis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic</td>
<td>9.19</td>
<td>.04</td>
<td>.94</td>
<td>.48</td>
</tr>
<tr>
<td>Thesis/Purpose</td>
<td>9.61</td>
<td>.04</td>
<td>1.1</td>
<td>.48</td>
</tr>
<tr>
<td>Supporting Material</td>
<td>9.66</td>
<td>.04</td>
<td>0.6</td>
<td>.45</td>
</tr>
<tr>
<td>Language</td>
<td>9.97</td>
<td>.04</td>
<td>0.8</td>
<td>.43</td>
</tr>
<tr>
<td>Organization</td>
<td>10.15</td>
<td>.04</td>
<td>1.0</td>
<td>.46</td>
</tr>
<tr>
<td>Articulation</td>
<td>10.28</td>
<td>.04</td>
<td>0.9</td>
<td>.44</td>
</tr>
<tr>
<td>Vocal Variety</td>
<td>10.46</td>
<td>.04</td>
<td>1.1</td>
<td>.45</td>
</tr>
<tr>
<td>Physical Behavior</td>
<td>10.67</td>
<td>.04</td>
<td>1.2</td>
<td>.45</td>
</tr>
</tbody>
</table>

### Rating Scale Structure

The rating Scale is divided into three classes: unsatisfactory, satisfactory, and excellent. The criteria are defined for each category, and the raters assign a number to each speaker’s competencies. Raters often have a tendency to group scores around the middle of the Scale values. McCroskey, Arnold and Pritchard (1967) established that the end points on a semantic differential were further from the points next to them than the other points were from each other. Some raters do not like to make extreme judgments. The Rasch model calibrates the nonlinearity of the rating Scale, thus eliminating this concern.

The Scale structure line shows that the rating Scale itself is appropriately structured – the categories are in pronounced locations and define different levels. (One could conceive of them as definite steps on a ladder.) The nine-point Scale provides more information than the three-point Scale. There is more precise measurement and greater ability to discriminate a speaker’s competencies. However, a nine-point Scale (low, medium and high categories in each class of unsatisfactory, satisfactory and excellent) is cumbersome and provides too many choices for the novice. Optimum information is provided by the six-point Scale, with a low and high category for each class. The three-point Scale has the benefit of being quick, clear, and easy to use. It is employed most often, and is an effective means of rating speeches.

### Speakers

It is possible to examine the differences in speakers’ competence if the speeches spread out along the line of inquiry. The **Competent Speaker Speech Evaluation Form** is a reliable test of separating speakers into strata of ability, and it provides accurate measurement. Because competence is reported in common units of measurement, comparisons of more or less ability among speakers are feasible.

In the present analyses, most speakers performed as expected. That is, the fit statistics do not indicate many speakers who have unexpected patterns of competency.

### Raters

Research conducted by members of the MESA Psychometric Laboratory at the University of Chicago, and many other places over the past 20 years reveals that judges, no matter how well trained, do not rate alike. In fact, it is not even desirable to attempt to force judges into one common mode. As we know from communication theory, every person has his or her own perceptual world and attends to different details. Bock and Bock (1981) discuss four general types of speech raters. In a Rasch analysis we assume each rater’s individuality and are not concerned with inter-rater reliability as an end to itself, which is only one of many indicators. Rather, it is the consistency with which each rater uses the evaluation form that is important. A Rasch analysis will adjust for the differences in type and severity of raters as long as they share a common understanding of the evaluation form (inter-rater agreement) and are individually consistent in their use of the rating Scale.
Inter-rater agreement is investigated. Each data analysis reveals the raters are observed to be in agreement more often than the expected percentage. This demonstrates a common understanding of The Competent Speaker Speech Evaluation Form across raters and groups of raters.

The raters spread out from very easy to very tough. A comparison of various groups shows there is no apparent bias to the instrument. No systematic, statistically significant differences are found in the use of The Competent Speaker Evaluation Form based on gender, ethnicity, teacher, school, or year.

All the items maintain their placement and still fit on the line of inquiry, which means the items are not being used differently by different groups. The speeches also maintain their order and fit. Thus, use of The Competent Speaker Speech Evaluation Form is consistent. Only a small percentage (approximately 8%) of raters either mis-fit, which means they are unpredictable in their evaluations; or over-fit, which means they are too consistent and do not discriminate in their evaluations. All the rest of the raters were internally consistent.

### The Competent Speaker Facets Map

The map below is a depiction of all the elements in the evaluation of competency in public speaking. It represents the combined data analysis.

On the right side of the map is the three-point rating Scale. The left side of the map shows the standardized units, or ruler by which we measure speaker competence, rater severity, and item difficulty.

The items, or competencies, cover a range of one and a half logits, or 150 points. The easiest competency (9.19) to achieve is *Chooses and narrows a topic appropriate to the audience and occasion*. The most difficult competency (10.67) to master is *Uses physical behaviors that support the verbal message*. The summary statistics for items are: measurement error .04; adjusted standard deviation .46; separation 10.84; and separation reliability .99.

The Competent Speaker Speech Evaluation Form creates a ruler that is sufficient to measure speeches covering a very wide range of ability – eight logits, or 800 points. The ruler discriminates fine differences all along the continuum of Public Speaking Competence. The most outstanding speeches have high of measures over 13.00; the poorest have low measures of less than 8.00. The distribution is normal. Summary statistics for speeches are: measurement error .19; adjusted standard deviation 1.10; separation 5.73; and separation reliability .97.

The raters also demonstrate a normal distribution and a very wide range of severity. The most severe raters have measures near 12.00; the easiest are below 8.00. The summary statistics for raters are: measurement error .32; adjusted standard deviation .77; separation 2.41; and separation reliability .85.

The Rasch model accounts for the difficulty of the item and the severity of the rater, which yields a much more precise measure of ability than mere average scores. This allows us to delve into the data in new ways to answer new questions.

### Conclusions of Rasch Analysis

The Competent Speaker Speech Evaluation Form is a viable instrument for assessing public speaking competence. The items cover the range of the variable and are all along the line of inquiry. Diverse raters use the form effectively, and diverse speakers demonstrate various levels of ability. Professors, community college instructors, graduate assistants, undergraduate students, and independent raters successfully use The Competent Speaker Speech Evaluation Form.

Based on these Rasch analyses, The Competent Speaker Speech Evaluation Form is found appropriate for general application. It provides a means for reliable assessment of Competent Public Speaking and is recommended for national distribution and use.
### Figure 1: Competent Speaker Facets Map

<table>
<thead>
<tr>
<th>Measr</th>
<th>Speaker</th>
<th>-Rater</th>
<th>-Items</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>+14</td>
<td>+ MORE</td>
<td>+</td>
<td>+ (3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ABILITY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+13</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+ ---</td>
</tr>
<tr>
<td>*</td>
<td>SEVERE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>RATER</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+12</td>
<td>+ ***.</td>
<td>+</td>
<td>+</td>
<td>+ ---</td>
</tr>
<tr>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*****</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*****</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*****</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*****</td>
<td>*</td>
<td>HARDER ITEM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*****</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+11</td>
<td>+ ***.</td>
<td>+ **</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>**</td>
<td>****</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>***</td>
<td>Nonverbal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*****</td>
<td>*****</td>
<td>Vocal Variety</td>
<td></td>
<td></td>
</tr>
<tr>
<td>****</td>
<td>Articulation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>****</td>
<td>**</td>
<td>Organization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+10</td>
<td>*</td>
<td>+ *****</td>
<td>+ Language</td>
<td>+ 2</td>
</tr>
<tr>
<td>**</td>
<td>*****</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*****</td>
<td>**.</td>
<td>Purpose Supporting Material</td>
<td></td>
<td></td>
</tr>
<tr>
<td>****</td>
<td>***.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>**</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>***</td>
<td>**</td>
<td>Topic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+9</td>
<td>+ **</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>*</td>
<td>EASIER ITEM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td>.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+8</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+ ---</td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+7</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LESS</td>
<td>EASY</td>
<td></td>
<td>ABILITY</td>
<td>RATER</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+6</td>
<td>+</td>
<td>+</td>
<td>+ (1)</td>
<td>+</td>
</tr>
</tbody>
</table>
APPENDIX D

THE NATIONAL COMMUNICATION ASSOCIATION’S CRITERIA FOR THE ASSESSMENT OF ORAL COMMUNICATION

A National Context

Assessment has received increasing attention throughout the 1970s and into the 1990s. Initially appearing in the standards developed by state departments of education, by 1980 over half of the states had adopted statewide student-testing programs. In Educational Standards in the 50 States: 1990, the Educational Testing Service reported that by 1985, over 40 states had adopted such programs, and between 1985 and 1990, an additional five states initiated statewide student-testing programs, bringing the number of such program to 47.

During the 1970s and 1980s, the number of different subjects and skills tested has also consistently increased, with additional attention devoted to how assessments are executed. Moreover, during this period, organizations, such as the National Assessment of Educational Progress, intensified and expanded the scope of their assessment procedures as well as extensively publicized the results of their findings nationally and annually.

By the end of 1989, the public recognized the significance of national educational assessments. In the Phi Delta Kappan-Gallup poll reported in the September 1989 issue of Phi Delta Kappan, 77 percent of the respondents favored “requiring the public schools in this community to use standardized national testing programs to measure academic achievement of students,” and 70 percent favored “requiring the public schools in this community to conform to national achievement standards and goals.”

Likewise, towards the end of the 1980s, colleges and universities began to realize that formal assessment issues were to affect them. For example, in its 1989-1990 Criteria for Accreditation, the Southern Association of Colleges and Schools—which provides institutional certification for over 800 colleges and universities in the South—held that “complete requirements for an associate or baccalaureate degree must include competence in reading, writing, oral communications and fundamental mathematical skills.” They also held that the general education core of colleges and universities “must provide components designed to ensure competence in reading, writing, oral communication and fundamental mathematical skills.”

In 1990, a series of reports appeared which suggested that systematic and comprehensive assessment should become a national educational objective. In February 1990, for example, the National Governors’ Association, in the context of President George H.W. Bush’s set of six educational goals, argued that, “National education goals will be meaningless unless progress toward meeting them is measured accurately and adequately, and reported to the American people.” The nation’s governors argued that “doing a good job of assessment” requires that “what students need to know must be defined,” “it must be determined whether they know it,” and “measurements must be accurate, comparable, appropriate, and constructive.” In July 1990, President Bush reinforced this line of reasoning in The National Education Goals: A Report to the Nation’s Governors. And, in September 1990, the National Governors Association extended and elaborated its commitment to assessment in Educating America: State Strategies for Achieving the National Education Goals: Report of the Task Force on Education.

Additionally, in 1990, in their report From Gatekeeper to Gateway: Transforming Testing in America, the National Commission on Testing and Public Policy recommended eight standards for assessment, arguing for more humane and multicultural assessment systems. Among other considerations, they particularly maintained that “testing policies and practices must be reoriented to promote the development of all human talent,” that “test scores should be used only when they differentiate on the basis of characteristics relevant to the opportunities being allocated, and that “the more test scores disproportionately deny opportunities to minorities, the greater the need to show that the tests measure characteristics relevant to the opportunities being allocated.”
NCA’s Assessment Activities

The evaluation and assessment of public address has been of central concern to the discipline of communication since its inception and to the National Communication Association when it was organized in 1914. In 1970, NCA formalized its commitment to assessment when it created the Committee on Assessment and Testing (now known by the acronym CAT) for “NCA members interested in gathering, analyzing and disseminating information about the testing of speech communication skills.” CAT has been one of the most active, consistent, and productive of NCA’s various committees and task forces.

Under the guidance of CAT, NCA has published several volumes exploring formal methods for assessing oral communication. These publications began to appear in the 1970s and have continued into the 1990s. In 1978, for example, the National Communication Association published Assessing Functional Communication, which was followed in 1984 by two other major publications, Large Scale Assessment of Oral Communication Skills: Kindergarten through Grade 12 and Oral Communication Assessment Procedures and Instrument Development in Higher Education.

In 1979, in Standards for Effective Oral Communication Programs, NCA adopted its first set of “standards” for “assessment and evaluation.” The first standards called for “school-wide assessment of speaking and listening needs of students,” “qualified personnel” to “utilize appropriate evaluation tools,” a “variety of data” and “instruments” which “encourage” “students’ desire to communicate.”

In 1986, in Criteria for Evaluating Instruments and Procedures for Assessing Speaking and Listening, NCA adopted an additional 15 “content” and “technical considerations” dealing “primarily with the substance of speaking and listening instruments” and “matters such as reliability, validity and information on administration.” These criteria included the importance of focusing on “demonstrated” speaking skills rather than “reading and writing ability,” adopting “assessment instruments and procedures” which are “free of sexual, cultural, racial, and ethnic content and/or stereotyping,” employing “familiar situations” which are “important for various communication settings” in test questions, using instruments which “permit a range of acceptable responses” and generate “reliable” outcomes, employing assessments which are consistent with other “results” and have “content validity,” and employing “standardized” procedures which “approximate the recognized stress level of oral communication” which are also “practical in terms of cost and time” and “suitable for the developmental level of the individual being tested.”

In 1987, at the NCA Wingspread Conference, “conference participants recommended that the chosen instrument conform to NCA guidelines for assessment instrument,” and they specifically suggested that “strategies for assessing speaking skills” should be directly linked to the content of oral communication performances and student speaking competencies. Prescribed communication practices were to determine the choice of assessment strategies, with the following content standards guiding formal evaluations: “determine the purpose of oral discourse;” “choose a topic and restrict it according to the purpose and the audience;” “fulfill the purpose” by “formulating a thesis statement,” “providing adequate support material,” “selecting a suitable organization,” “demonstrating careful choice of words,” “providing effective transitions,” “demonstrating suitable inter-personal skills;” employing “vocal-variety in rate, pitch, and intensity;” “articulate clearly;” “employ the level of American English appropriate to the designated audience;” and “demonstrate nonverbal behavior that supports the verbal message.” Additionally, the Wingspread Conference participants considered strategies for assessing listening and for training assessors [see: Communication Is Life: Essential College Sophomore Speaking and Listening Competencies (Washington, D.C.: National Communication Association, 1990, pp. 51-74).

In July 1990, a National Conference on Assessment was sponsored by NCA, the NCA Committee on Assessment and Testing or CAT, and the NCA Educational Policies Board (EPB). The conference generated several resolutions regarding assessment.* Some of these resolutions reaffirm existing NCA oral communication assessment policies. Others provide criteria for resolving new issues in assessment. Still others seek to integrate and establish a more coherent relationship among the criteria governing oral communication assessment. The recommended assessment criteria are detailed on the next page.

**General Criteria**

1. Assessment of oral communication should view competence in oral communication as a gestalt of several interacting dimensions. At a minimum, all assessments of oral communication should include an assessment of knowledge (understanding communication process, comprehension of the elements, rules, and dynamics of a communication event, awareness of what is appropriate in a communication situation), an assessment of skills (the possession of a repertoire of skills and the actual performance of skills), and an evaluation of the individual’s attitude toward communication (e.g., value placed on oral communication, apprehension, reticence, willingness to communicate, readiness to communicate).

2. Because oral communication is an interactive and social process, assessment should consider the judgment of a trained assessor as well as the impressions of others involved in the communication act (audience, interviewer, other group members, conversant), and may include the self report of the individual being assessed.

3. Assessment of oral communication should clearly distinguish speaking and listening from reading and writing. While some parts of the assessment process may include reading and writing, a major portion of the assessment of oral communication should require speaking and listening. Directions from the assessor and responses by the individual being assessed should be in the oral/aural mode.

4. Assessment of oral communication should be sensitive to the effects of relevant physical and psychological disabilities on the assessment of competence. (e.g., with appropriate aids in signal reception, a hearing impaired person can be a competent empathic listener.)

5. Assessment of oral communication should be based in art on atomistic/analytic data collected and on a holistic impression.

**Criteria for the Content of Assessment**

6. Assessment of oral communication for all students should include assessment of both verbal and non-verbal aspects of communication and should consider competence in more than one communication setting. As a minimum assessment should occur in the one-to-many setting (e.g. public speaking, practical small group discussion) and in the one-to-one setting (e.g., interviews, interpersonal relations).

7. Assessment of speech majors and other oral communication specialists could include in addition assessment in specialized fields appropriate to the course of study followed or the specialty of the person being assessed.

**Criteria for Assessment Instruments**

8. The method of assessment should be consistent with the dimension of oral communication being assessed. While knowledge and attitude may be assessed in part through paper and pencil instruments, speaking and listening skills must be assessed through actual performance in social settings (speaking before an audi-
ence, undergoing an interview, participating in a group discussion, etc.) appropriate to the skill(s) being assessed.

9. Instruments for assessing oral communication should describe degrees of competence. Either/or descriptions such as “competent” or “incompetent” should be avoided as should attempts to diagnose reasons why individuals demonstrate or fail to demonstrate particular degrees of competence.

10. Instruments for assessing each dimension of oral communication competence should clearly identify the range of responses which constitute various degrees of competence. Examples of such responses should be provided as anchors.

11. Assessment instruments should have an acceptable level of reliability, e.g. test/retest reliability, split-half reliability, alternative forms reliability, inter-rater reliability, and internal consistency.

12. Assessment instruments should have appropriate validity: content validity, predictive validity, and concurrent validity.

13. Assessment instruments must meet acceptable standards for freedom from cultural, sexual, ethical, racial, age, and developmental bias.

14. Assessment instruments should be suitable for the developmental level of the individual being assessed.

15. Assessment instruments should be standardized and detailed enough so that individual responses will not be affected by an administrator’s skill in administering the procedures.

Criteria for Assessment Procedures and Administration

16. Assessment procedures should protect the rights of those being assessed in the following ways: administration of assessment instruments and assessment and the uses of assessment results should be kept confidential and be released only to an appropriate institutional office, to the individual assessed, or if a minor, to his or her parent or legal guardian.

17. Use of competence assessment as a basis for procedural decisions concerning an individual should, when feasible, be based on multiple sources of information, including especially a) direct evidence of actual communication performance in school and/or other contexts, b) results of formal competence assessment, and c) measures of individual attitudes toward communication (e.g., value placed on oral communication, apprehension, reticence, willingness to communicate, and readiness to communicate).

18. Individuals administering assessment procedures for oral communication should have received sufficient training by speech communication professionals to make their assessment reliable. Scoring of some standardized assessment instruments in speaking and listening may require specialized training in oral communication on the part of the assessor.

Criteria for Assessment Frequency

Periodic assessment of oral communication competency should occur annually during the educational careers of students. An effective systematic assessment program minimally should occur at educational levels K, 4, 8, 12, 14, and 16.

Criteria for the Use of Assessment Results

The results of student oral communication competency assessment should be used in an ethical, non-discriminatory manner for such purposes as:

19. Diagnosing student strengths and weaknesses;

20. Planning instructional strategies to address student strengths and weaknesses;

21. Certification of student readiness for entry into and exit from programs and institutions;

22. Evaluating and describing overall student achievement;
23. Screening students for programs designed for special populations;
24. Counseling students for academic and career options; and
25. Evaluating the effectiveness of instructional programs.

No single assessment instrument is likely to support all these purposes. Moreover, instruments appropriate to various or multiple purposes typically vary in length, breadth/depth of content, technical rigor, and format.
APPENDIX E

COMPLIANCE BY “THE COMPETENT SPEAKER” WITH THE “NCA CRITERIA FOR THE ASSESSMENT OF ORAL COMMUNICATION”

The left column contains abbreviated descriptions of the criteria contained in the “NCA Criteria for the Assessment of Oral Communication.” The right column identifies which criteria are met by The Competent Speaker assessment instrument.

<table>
<thead>
<tr>
<th>Notation of Compliance by General Criteria</th>
<th>The Competent Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Assesses knowledge, skills &amp; attitude</td>
<td>Public speaking skills only</td>
</tr>
<tr>
<td>2. Judged by trained assessor</td>
<td>Yes (via manual and training tape)</td>
</tr>
<tr>
<td>3. Assesses speaking &amp; listening</td>
<td>Speaking only</td>
</tr>
<tr>
<td>4. Sensitive to assessee’s disabilities</td>
<td>Yes (via manual and training tape)</td>
</tr>
<tr>
<td>5. Assessment based on atomistic data &amp; holistic impression</td>
<td>Designed for atomistic eval.; holistic possible</td>
</tr>
</tbody>
</table>

Criteria for Content of Assessment

1. Assesses verbal & nonverbal in multiple settings | Verbal & nonverbal in public speaking setting |
2. Assesses specialty area of speech majors | Not applicable |

Criteria for Assessment Instruments

1. Skills assessed by performance in public speaking setting | Yes |
2. Assesses degree of competence | Yes |
3. Identifies range of responses which constitutes degrees of competence | Yes |
4. Meets acceptable level of reliability | Yes |
5. Meets acceptable standards of validity | Yes |
6. Free of cultural, sexual, ethnic, racial, age & developmental bias | Yes |
7. Suitable for developmental level of assessee | Yes (college sophomore/Grade 14 level) |
8. Standardized procedures for administering | Yes |

Criteria for Assessment Procedures & Administration

1. Procedures protect rights of assessee | Yes (procedures outlined in manual) |
2. Assessment for procedural decision based on multiple sources of info. | Use for procedural decisions in manual |
3. Assessors trained by speech and communication professionals | Yes (via manual training tape) |
VIII. Bibliography


Wright, B. D. & Stone, M. H. Best test design. Chicago, IL: MESA Press.
