

# AN INSTRUMENT FOR MEASURING THE SOURCE CREDIBILITY OF BASIC SPEECH COMMUNICATION INSTRUCTORS

James C. McCroskey, William Holdridge, and J. Kevin Toomb

A NUMBER of studies have been reported which have investigated the dimensionality of the source credibility construct and provided suggested scales for the measurement of credibility.<sup>1</sup> Most of these studies have focused on the credibility of public figures and have used college students as subjects. None have specifically looked at teachers. Tucker<sup>2</sup> has noted the error in assuming that these scales can be used for other types of sources (such as teachers). As Tucker points out, varying the subject-type or source-type may cause the dimensionality of source credibility to change. Students may not respond to teacher-sources on the same dimensions on which they respond to public figures and, even if the same dimensions of re-

sponse are present, the best scales for measuring those responses may vary. Recent research on types of sources other than teachers indicates that the caution requested by Tucker is appropriate.<sup>3</sup> The dimensionality of source credibility can fluctuate from one type of source to another, and even in regard to the same dimension (competence, for example), the scales that best measure that dimension may not be the same for two different source-types.

It is essential, therefore, that if we wish to measure teacher credibility, an instrument must be developed specifically for that purpose. The current research was designed to achieve that end.

## METHOD

*Scales.* This investigation employed 46 semantic differential-type scales representing the dimensions of source credibility observed in the previously cited research by Norman, McCroskey, Mark-

Mr. McCroskey is professor and chairman of the Department of Speech Communication at West Virginia University. Mr. Holdridge is a doctoral candidate in Speech Communication at the University of Illinois. Mr. Toomb is a doctoral candidate in Communication at Florida State University.

<sup>1</sup> W. T. Norman, "Toward an Adequate Taxonomy of Personality Attributes: Replicated Factor Structure in Peer Nomination Personality Ratings," *Journal of Abnormal and Social Psychology*, 66 (June 1963), 574-583; J. C. McCroskey, "Scales for the Measurement of Ethos," *Speech Monographs*, 33 (March 1966), 65-72; D. Markham, "The Dimensions of Source Credibility of Television Newscasters," *Journal of Communication*, 18 (March 1968), 57-64; J. L. Whitehead, Jr., "Factors of Source Credibility," *Quarterly Journal of Speech*, 54 (February 1968), 59-63; D. K. Berlo, J. B. Lemert, and R. Mertz, "Dimensions for Evaluating the Acceptance of Message Sources," *Public Opinion Quarterly*, 33 (Winter 1969), 563-576.

<sup>2</sup> R. K. Tucker, "On the McCroskey Scales for the Measurement of Ethos," *Central States Speech Journal*, 22 (Summer 1971), 127-129.

<sup>3</sup> J. C. McCroskey, M. D. Scott, and T. J. Young, "The Dimensions of Source Credibility for Spouses and Peers," paper presented at the Western Speech Communication Association convention, Fresno, 1971; J. C. McCroskey, T. Jensen, C. Todd, and J. K. Toomb, "Measurement of the Credibility of Organization Sources," paper presented at the Western Speech Communication convention, Honolulu, 1972; J. C. McCroskey, T. Jensen, and C. Todd, "The Generalizability of Source Credibility Scales for Public Figures," paper presented at the Speech Communication Association convention, Chicago, 1972; J. C. McCroskey, T. Jensen, and C. Valencia, "Measurement of the Credibility of Spouses and Peers," paper presented at the International Communication Association convention, Montreal, 1973.

ham, and Berlo, Lemert, and Mertz. In a pilot study which led to the current investigation, all of the scales with high loadings on factors from these studies were included, a total of 53 items. After the pilot phase of the study, which investigated four source-types, 11 items were omitted from the original item pool. The items omitted failed to have a satisfactory factor loading on any factor for any source-type.<sup>4</sup> Four additional items were added to the pool after the pilot study. These scales were believed to be related to factors observed in the pilot study which had only two or three items with satisfactory loadings and for which there were no additional items in the original pool that appeared to be related.<sup>5</sup>

*Teachers—Subjects.* The first sample in this study included 642 students in the basic course in speech communication at Illinois State University. At the time that the study was conducted, this course was taught by means of mass lectures with accompanying laboratory sections. There were nine senior-faculty lecturers in the class. Each student involved in the study responded to one (randomly determined) lecturer. The second sample in the study

also involved students in the basic speech communication course at Illinois State. The 663 subjects in this part of the study responded to their laboratory instructor. Thirty-seven sections of the course were employed, each with a different teacher. The third part of the study involved 575 students in the basic speech communication course at the University of Illinois. At the time that this study was conducted, each section of the course was taught independently, with the syllabus being prepared by the individual instructor. Thirty-five sections of the course, with 19 separate instructors, were involved. All three samples were tested during the fall semester, 1971-72.

*Data Analyses.* The data from the three samples were analyzed separately. The semantic differential data were submitted to principal component factor analyses and varimax rotation. Unity was inserted in the diagonals and an eigenvalue of 1.0 was established as the criterion for termination of factor extraction. For an item to be considered loaded on a resulting factor, a loading of .60 or higher was required with no loading of .40 or higher on any other factor.<sup>6</sup> For a factor to be considered meaningful, the requirement was that two scales must have satisfactory loadings on that factor.

All data analyses were performed with the cooperation of the computer centers at Illinois State University and West Virginia University.

<sup>4</sup> A satisfactory loading was considered to be .60 or higher on one factor with no secondary loading higher than .40.

<sup>5</sup> The scales employed were: intelligent-unintelligent, sociable-unsociable, nervous-poised, cheerful-gloomy, tense-relaxed, sinful-virtuous, believable-unbelievable, good-natured-irritable, intellectual-narrow, cooperative-negativistic, outgoing-withdrawn, dishonest-honest, meek-aggressive, valuable-worthless, selfish-unselfish, calm-anxious, inexperienced-experienced, verbal-quiet, logical-illogical, undependable-responsible, head-strong-mild, friendly-unfriendly, confident-lacks confidence, untrained-trained, unsympathetic-sympathetic, admirable-contemptible, awful-nice, qualified-unqualified, extroverted-introverted, just-unjust, unpleasant-pleasant, timid-bold, energetic-tired, good-bad, repulsive-attractive, uninformed-informed, composed-excitable, incompetent-competent, cruel-kind, talkative-silent, expert-inexpert, passive-active, impressive-unimpressive, adventurous-cautious, crude-refined, and reliable-unreliable.

<sup>6</sup> These loading criteria are relatively conservative and were chosen for that reason. In common research practice raw scores rather than factor scores are usually used from data generated by semantic differential-type instruments. Inclusion of items with less pure loadings results in correlated dimension scores even though the dimensions are the product of orthogonal factor analysis. More liberal criteria would indicate more items that could be used to measure obtained dimensions, but the resulting dimension scores would be increasingly interrelated. Use of such scores, consequently, would introduce systematic error into the research, and should be avoided.



## RESULTS

Factor analysis of the data provided by the subjects in the first sample (those responding to mass lecturers at Illinois State University) indicated the presence of five dimensions of response. These dimensions were labeled "Character," "Sociability," "Composure," "Extroversion," and "Competence." These factors accounted for 62 per cent of total variance of the scales with satisfactory loadings. (See Table 1).

Factor analysis of the data provided in the second sample (those responding to their laboratory instructors at Illinois State University) also yielded five factors which were labeled in the same manner as above. These five factors accounted for 60 per cent of the total variance of

the scales with satisfactory loadings. (See Table 2).

Factor analyses of the data from the students in the third sample (those responding to instructors teaching independent sections at the University of Illinois) resulted in a four-factor solution which accounted for 60 per cent of the total variance of the items with satisfactory loadings. These factors were labeled "Sociability-Character," "Composure," "Extroversion," and "Competence."

## DISCUSSION

The results from the three samples in this study suggest the presence of five dimensions of source credibility for teachers. For the first two samples, the factor analyses resulted in five clear dimensions

TABLE 1  
ROTATED FACTOR LOADINGS FOR MASS LECTURERS

Item	Factor				
	Character	Sociability	Composure	Extroversion	Competence
Intelligent-Unintelligent	.01	.15	-.11	-.17	.73
Sociable-Unsociable	.20	.61	-.16	-.22	.34
Nervous-Poised	-.13	-.09	.68	.36	-.13
Cheerful-Gloomy	.11	.68	-.13	-.30	.34
Tense-Relaxed	-.11	-.15	.70	.39	-.16
Sinful-Virtuous	-.66	.09	.11	-.01	-.12
Believable-Unbelievable	.19	.34	-.15	-.03	.69
Good-natured-Irritable	.24	.64	-.13	-.18	.39
Intellectual-Narrow	.12	.30	-.17	-.09	.73
Outgoing-Withdrawn	.15	.34	-.12	-.69	.21
Dishonest-Honest	-.65	-.01	.14	.08	-.34
Meek-Aggressive	-.10	-.01	.17	.65	.02
Valuable-Worthless	.30	.25	-.14	-.19	.71
Selfish-Unselfish	-.71	-.21	.01	.01	-.19
Calm-Anxious	.04	.13	-.77	-.19	.26
Verbal-Quiet	-.06	.11	-.14	-.73	.14
Logical-Illogical	.22	.25	-.20	-.16	.69
Confident-Lacks Confidence	.06	.10	-.39	-.65	.30
Qualified-Unqualified	.19	.10	-.15	-.28	.75
Extroverted-Introverted	-.02	.15	-.09	-.71	.18
Timid-Bold	-.11	-.03	.17	.65	.06
Energetic-Tired	.23	.19	.01	-.68	.29
Composed-Excitable	.08	.00	-.73	-.03	.19
Incompetent-Competent	-.32	-.13	.10	.18	-.64
Cruel-Kind	-.60	-.33	-.04	.04	-.32
Talkative-Silent	-.07	.12	-.08	-.75	.10
Expert-Inexpert	.22	.00	-.12	-.29	.77
Impressive-Unimpressive	.35	.21	-.18	-.27	.62
Adventurous-Cautious	.09	.02	-.10	-.69	.21
Reliable-Unreliable	.34	.16	-.09	-.12	.74
Eigenvalue	2.60	2.12	2.65	5.16	6.15
Variance	.09	.07	.09	.17	.20
Eigenvalue Total = 18.85				Variance Total = 62	

TABLE 2  
ROTATED FACTOR LOADINGS FOR SMALL SECTION INSTRUCTORS

Item	Factor				
	Character	Sociability	Composure	Extroversion	Competence
Intelligent-Unintelligent	.05	.39	-.21	-.09	-.61
Sociable-Unsociable	.26	.72	-.06	-.24	-.25
Nervous-Poised	-.23	-.10	.62	.24	.24
Cheerful-Gloomy	.16	.66	.01	-.22	-.22
Tense-Relaxed	-.27	-.06	.63	.33	.24
Good-natured-Irritable	.28	.73	-.09	-.16	-.26
Cooperative-Negativistic	.31	.62	-.18	-.04	-.34
Meek-Aggressive	-.22	-.01	.02	.68	.14
Selfish-Unselfish	.60	-.11	.14	.06	.26
Calm-Anxious	.11	.08	-.74	-.03	-.22
Inexperienced-Experienced	-.25	-.04	.16	.29	.66
Verbal-Quiet	-.01	.33	-.04	-.63	-.06
Friendly-Unfriendly	.39	.67	-.08	-.23	-.23
Untrained-Trained	-.30	-.08	.13	.24	.72
Unsympathetic-Sympathetic	-.66	-.25	.07	.02	.25
Awful-Nice	-.62	-.35	.08	.03	.32
Qualified-Unqualified	.24	.34	-.14	-.13	-.73
Timid-Bold	-.22	.01	.09	.74	.20
Composed-Excitable	-.05	.09	-.68	.14	-.03
Cruel-Kind	-.61	-.33	.09	.03	.35
Talkative-Silent	-.06	.32	.03	-.64	-.05
Expert-Inexpert	.03	.28	-.13	-.20	-.67
Reliable-Unreliable	.15	.36	-.08	-.06	-.67
Eigenvalue	2.45	3.22	2.02	2.41	3.69
Variance	.11	.14	.09	.10	.16
Eigenvalue Total = 13.96				Variance Total = 60	

that were highly similar. The results from the third sample, while indicating the presence of only four dimensions, replicated the earlier results for the dimensions of "Composure," "Extroversion," and "Competence." The only deviation indicated by the results from the third sample was the collapse of the "Character" and "Sociability" dimensions into one dimension. Since in most cases it is better to have too much information about teacher credibility than too little, it would seem best to consider these results as indicative of the presence of five dimensions of teacher credibility. In some cases it would appear that "Character" and "Sociability" can operate independently, while in other cases they may operate conjointly. The worst that can happen as a result of treating the two dimensions as separate if in fact they are not, is the obtaining of two highly correlated scores. However, if one were to treat the two dimensions as one

when in fact they were operating independently, the obtained score would be quite meaningless.

The instrument that is recommended for measuring teacher credibility appears in Table 4. However, if one were interested in including more scales for a dimension than are present among the recommended scales, additional scales could be selected from Tables 1-3 that would be appropriate for teachers under varying types of instructional patterns (mass lectures, laboratory sections, or independent sections).

When the instrument in Table 4 is used, the scales can be rearranged in random order, and the polarity determined randomly. The numbers in the spaces in Table 4 are included for illustration only, and should not appear in normal use. They indicate how a mark in a given space should be scored. Scales 1-2 measure "Competence," 3-6 measure "Extroversion," 7-8 measure "Character," 9-11



TABLE 3  
FACTOR LOADINGS FOR INDEPENDENT SECTIONS

Scales	Factors			
	Character-Sociability	Extroversion	Competence	Composure
Sociable-Unsociable	.71	.26	.05	-.15
Nervous-Poised	.25	.23	.13	-.73
Cheerful-Gloomy	.73	.17	.01	-.18
Tense-Relaxed	.36	.25	.10	-.71
Good-Natured-Irritable	.82	-.01	.07	-.17
Cooperative-Negativistic	.76	.05	.26	-.09
Meek-Aggressive	.06	-.67	.03	-.06
Valuable-Worthless	.67	.17	.37	-.12
Calm-Anxious	.17	.03	.18	.67
Verbal-Quiet	.18	.69	.04	-.07
Friendly-Unfriendly	.79	.16	.12	-.11
Unsympathetic-Sympathetic	-.68	-.04	-.33	-.07
Admirable-Contemptible	.70	.12	.35	-.18
Awful-Nice	-.75	.02	-.31	-.13
Just-Unjust	.68	.08	.37	-.10
Unpleasant-Pleasant	-.68	.13	-.33	-.12
Timid-Bold	.01	-.71	.21	.05
Talkative-Silent	.03	.65	.24	-.02
Expert-Inexpert	.14	-.39	.69	-.14
Crude-Refined	.21	.01	-.61	-.25
Reliable-Unreliable	.30	.10	.63	-.12
Eigenvalue	6.21	2.32	2.21	1.79
Variance	.30	.11	.11	.09
Total Eigenvalue = 12.53			Total Variance = .60	

TABLE 4  
RECOMMENDED TEACHER CREDIBILITY MEASURE

*Instructions:* The following are a series of attitude scales. You are asked to evaluate your instructor in terms of the adjectives on each scale. For example, if you think your instructor is very tall you might mark the following scale as below:

Tall : X : — : — : — : — : — : — : Short

Of course, if you consider your instructor to be shorter you would mark your "X" nearer the "short" adjective. The middle space on each scale should be considered "neutral." Mark this space if you feel neither adjective on the scale applies to your instructor or if you feel both apply equally.

*My Instructor in This Course*

1.	Expert	: 7	: 6	: 5	: 4	: 3	: 2	: 1	: Inexpert
2.	Unreliable	: 1	: 2	: 3	: 4	: 5	: 6	: 7	: Reliable
3.	Meek	: 1	: 2	: 3	: 4	: 5	: 6	: 7	: Aggressive
4.	Verbal	: 7	: 6	: 5	: 4	: 3	: 2	: 1	: Quiet
5.	Bold	: 7	: 6	: 5	: 4	: 3	: 2	: 1	: Timid
6.	Silent	: 1	: 2	: 3	: 4	: 5	: 6	: 7	: Talkative
7.	Unselfish	: 7	: 6	: 5	: 4	: 3	: 2	: 1	: Selfish
8.	Kind	: 7	: 6	: 5	: 4	: 3	: 2	: 1	: Cruel
9.	Poised	: 7	: 6	: 5	: 4	: 3	: 2	: 1	: Nervous
10.	Tense	: 1	: 2	: 3	: 4	: 5	: 6	: 7	: Relaxed
11.	Anxious	: 1	: 2	: 3	: 4	: 5	: 6	: 7	: Calm
12.	Unsociable	: 1	: 2	: 3	: 4	: 5	: 6	: 7	: Sociable
13.	Cheerful	: 7	: 6	: 5	: 4	: 3	: 2	: 1	: Gloomy
14.	Irritable	: 1	: 2	: 3	: 4	: 5	: 6	: 7	: Good-natured

measure "Composure," and 12-14 measure "Sociability."

*Evaluation of the Measure*

Evaluation of the suggested measure of

teacher credibility should be based on three criteria: the reliability of the instrument, the validity of the instrument as a measure of teacher credibility, and the ability of the instrument to predict

student learning. We shall consider each of these in turn.

*Reliability.* The reliability of this instrument was tested in a follow-up study. During the spring semester, 1971-72, 948 students completed the instrument reported in Table 4. Each student responded to one section instructor in the basic speech communication course at Illinois State University. Thirty-six different instructors were evaluated. Internal reliability estimates (based on the Hoyt procedure<sup>7</sup>) for the five dimensions of response ranged from a low of .91 for "Competence" to a high of .94 for "Extroversion." Forty-six students were involved in a test-retest situation with a two week interval between the testing sessions. Reliability estimates for the five dimensions ranged from .82 for "Competence" to .86 for "Sociability." These reliability estimates are well within the range normally considered satisfactory.

*Construct Validity.* Since there was no established measure of teacher credibility available against which to compare the present measure, no statistical estimate of the validity of the instrument as a measure of teacher credibility was possible. Consequently, the construct or face validity of the measure must be considered. Since the pool of items upon which the instrument was built represented a wide variety of previously used scales for source credibility, there is reason to believe that the item pool, and thus the resulting factors and scales, are representative of the credibility construct. In addition, a subjective examination of the resulting factors and chosen scales suggests face validity. All of the dimensions appear related to credibility, and each of the scales seems to be logically as-

sociated with the factor on which it was highly loaded.

*Predictive Validity.* Since the major purpose underlying the present study was the development of a valid means of teacher evaluation based on student response, the most important test of the validity of the resulting instrument is whether it is a significant predictor of student learning. This question has been subjected to two direct tests.

The first test was based on the assumption that a student's willingness to sign up for another course from an instructor and/or his willingness to recommend the course to a friend are related to information acquisition.<sup>8</sup> The subjects in the two primary Illinois State samples, those who responded to mass lecturers and those who responded to laboratory instructors, were asked two additional questions designed to determine how likely they were to expose themselves voluntarily again to the instructor and whether they would recommend that a friend of theirs do so. The questions asked were: "If you had room in your schedule for an elective course, how likely would it be that you would sign up for

<sup>8</sup> The rationale for this assumption is two-fold. First, no information can be acquired by a student from a teacher unless the student is willing to expose him or herself to communication from that teacher. Mere exposure, of course, will not guarantee that learning will occur. However, non-exposure will guarantee non-learning. Considerable research has indicated the existence of the selective exposure phenomenon (see E. Katz, "On Reopening the Question of Selectivity in Exposure to Mass Communication," in R. P. Abelson et al., *Theories of Cognitive Consistency: A Sourcebook*, Chicago: Rand McNally, 1968). Second, although most of the research in this area has focused on receivers' attitudes on message topics as causal agents, recent research has indicated that a highly reliable predictor of selective exposure is source credibility (see L. R. Wheeless, "The Effects of Attitude, Credibility, and Homophily on Selective Exposure to Information," paper presented at the International Communication Association convention, Montreal, 1973). While the present research focused on projected future exposure, later research should consider direct effects on present exposure, such as class attendance rates.

<sup>7</sup> C. Hoyt, "Test Reliability Estimated by Analysis of Variance," *Psychometrika*, 6 (June 1941), 153-160.



another course with this instructor?" and "Would you suggest that a friend of yours sign up for a course from this instructor?" The subjects were asked to respond to an eleven-step continuum bound by "very likely" and "very unlikely."

Table 5 reports the observed correlations between the credibility dimensions and these indications of projected exposure and recommended exposure. All of the observed correlations were statistically significant ( $p < .01$ ). High scores on each dimension were associated with greater likelihood of exposure. Multiple correlation analyses indicated that from 45 to 58 per cent of the variance in projected exposure could be predicted by credibility. "Competence" and "Sociability" were consistently the best predictors. These results, therefore, indicate the validity of the credibility measure developed in the present study for the prediction of probable information acquisition under exposure conditions permitting voluntary choice. Students indicate a desire to take courses from teachers perceived as credible, as measured by the present instrument, and to avoid courses from teachers perceived as less credible.

The second test of the predictive validity of the instrument was related to information acquisition under conditions

of non-voluntary exposure. During the fall semester, 1972-73, 118 students in the basic course in speech communication at West Virginia University completed the teacher-credibility instrument in reference to their course instructors.<sup>9</sup> Ten instructors were involved. The students were exposed to an experimental message by their respective instructors in the context of a regular class period. The message was related to the course, and the information included was not available from any other source. The students were not forewarned that they would be tested over the material. However, their immediate recall was measured by a cloze-procedure test<sup>10</sup> immediately after exposure to the message.

Analysis of the results indicated that the only credibility dimension significantly correlated to immediate recall was "Competence." These results, then, provide only marginal support for the predictive validity of the instrument. It should be noted, however, that there was

<sup>9</sup> The results reported here are part of a broader study. For a report of the complete study, see L. R. Wheelless, "The Relationship of Course Attitudes, Instructor Credibility, Attraction, and Homophily to Immediate Recall and Student-Instructor Interaction," paper presented at the Speech Communication Association convention, New York, 1973.

<sup>10</sup> This test included 54 deleted words. The split-half reliability of the test (corrected) was .92.

TABLE 5  
CORRELATIONS BETWEEN CREDIBILITY DIMENSIONS  
AND VOLUNTARY EXPOSURE AND RECOMMENDED EXPOSURE

Credibility Dimension	Voluntary Exposure (Take Another Course)		Recommended Exposure (Friend Take Course)	
	Mass Lecturers	Laboratory Instructors	Mass Lecturers	Laboratory Instructors
Competence	.62	.55	.71	.51
Extroversion	.40	.31	.36	.27
Composure	.37	.34	.38	.33
Character	.44	.48	.29	.50
Sociability	.62	.71	.60	.53
Multiple Correlation	.73	.76	.76	.67
Per cent of Variance Predictable	53	58	58	45

very little variance found in the credibility scores observed in this study. Essentially, each teacher was perceived as highly credible on each dimension by almost all of the students. This lack of variance could account for the low and nonsignificant correlations observed. In addition, of the approximately 200 students who should have been present for the study, only 118 were actually in attendance. The missing 40 per cent may have seen the instructors as much less credible and thus may have chosen not to expose themselves to the instructor that day. This explanation, of course, is speculative, but the results discussed above concerning credibility and exposure would indicate the explanation may have merit. Nevertheless, at this point only marginal support is available

for the predictive validity of the instrument with respect to immediate recall of information.

### *Conclusions*

The results of this investigation indicate that the teacher-credibility instrument that was developed is a reliable measure, has satisfactory construct and face validity, and has predictive validity at least for projected future exposure. The instrument is potentially useful to the speech communication instructor for purposes of teacher evaluation when standardized, criterion based measures of student learning are not feasible.<sup>11</sup>

<sup>11</sup> For an excellent discussion of the desirability of using student-generated teacher evaluations, see C. N. Wise, "Student Ratings of Teachers: A Perspective for Speech Communication," *Western Speech*, 37 (Fall 1973), 196-203.