

Test of Oral Communication Skills, version 2 (TOCS-2)

Test Manual



Madison Assessment



THE CENTER FOR
**ASSESSMENT &
RESEARCH**
STUDIES

Laura M Williams & S. Jeanne Horst & Donna L. Sundre
2014

MSC 6806
Harrisonburg, VA 22807
540.568.6706 Phone
540.568.7878 Fax
assessment@jmu.edu
www.jmu.edu/assessment

Table of Contents
Test of Oral Communication Skills, version 2 (TOCS-2)

Section 1. The Nature of the Instrument	3
Section 2. Intended Use.....	3
2.1. <i>Appropriate and inappropriate uses and interpretations</i>	3
2.2. <i>Target population</i>	3
2.3. <i>Qualifications of users</i>	3
Section 3. Test Development	4
3.1. <i>Academic and theoretical basis</i>	4
3.2. <i>Item type selection</i>	4
3.3. <i>Item pool and scale development process</i>	4
Table 1	5
3.4. <i>Subscores and their development</i>	5
Section 4. Administrative Procedures	6
4.1. <i>Proctor qualifications and training.....</i>	6
4.2. <i>Testing procedures.....</i>	6
4.2.1. <i>Computer-based administration.....</i>	6
Section 5. Technical Information.....	6
5.1. <i>Scoring and interpretation.....</i>	6
5.1.1. <i>Standard Setting.....</i>	7
5.2. <i>Evidence of reliability.....</i>	7
Table 2.....	7
5.3. <i>Evidence of validity.....</i>	8
Table 3.....	8
Table 4.....	9
5.4. <i>Norming.....</i>	9
Table 5.....	9
Section 6. Additional Information	10
6.1. <i>Future work to be conducted</i>	10
6.2. <i>Where to find additional information</i>	10
Section 7. References	10

The TOCS-2 Test Manual

Section 1. The Nature of the Instrument

The Test of Oral Communication Skills II (TOCS-2) is an assessment package designed to measure undergraduate general education competence in human communication (100-item cognitive test) and willingness and confidence to engage in oral communication (11-item Attitudes Toward Communications; ATC; Ball, Jurich & Anderson, 2014). The test was written to align with the National Communication Association (NCA) standards (1998). The ATC items measure affective components of communication thought to influence student deployment of communication competencies: willingness and confidence. We believe that inclusion of these affective measures are an important supplement to the assessment of cognitive and communication competencies. The TOCS-2 is administered via computer and includes innovative visual and audio media clips. Versions of the test appropriate for visually and/or auditory-impaired students are available.

Section 2. Intended Use

2.1. Appropriate and inappropriate uses and interpretations

The TOCS-2 was created to provide information about the effects of curriculum and instruction on students' learning. It was designed to measure students' knowledge of core communication constructs and ability to make educated decisions about oral communication behaviors. TOCS-2 was developed for use at the programmatic level; information provided in this manual is based upon data gathered from use of the TOCS-2 as a graded final examination. Therefore, test results may be used to inform curriculum and instructional improvements at the program or institution level as well as student-level evaluation. Test users are responsible for collecting validity evidence for any uses of the test other than those recommended here (AERA, APA, & NCME, 2000).

2.2. Target population

The TOCS-2 is intended to measure learning of key communication concepts for undergraduate college students. The instrument was designed to be appropriate for students in any college-level general education oral communication curriculum. At the institution in which the instrument was developed, all first-year students complete one of three communication courses that address the objectives described in Section 3.1. Although the TOCS-2 was designed with all undergraduate college students in mind, the normative data provided cannot be representative of all other institutions. Thus, other institutions will want to review descriptions of the norming groups that framed the writing of this test manual. Please refer to Section 5.4, *Norming*, for more information about these groups. Other institutions are encouraged to gather their own norming information and to explore their own important validity questions.

2.3. Qualifications of users

Test users must be trained to administer assessments using computers as students will be asked to respond to items that include audio and video playback, , and traditional multiple-choice selected-response items. TOCS-2 is not available in a paper-and-pencil version, but closed captioning and detailed descriptions of visual material are available for test takers who are visually or auditorily impaired. The *Proctor qualifications and training section* of this manual (Section 4.1) provides more information about how proctors can be trained to conduct computerized test administration. Madison Assessment, LLC will complete scoring¹ for this instrument. In addition, test users should be knowledgeable about how to interpret test results.

¹ Scoring (correct/incorrect) includes all instrument items and, if desired, up to 20 additional items added to the instrument by those institutions seeking to further evaluate the construct.

Section 3. Test Development

3.1. Academic and theoretical basis

The TOCS-2 was designed to evaluate student learning in four general education objectives related to oral communication competence. The objectives, listed below, align with NCA (1998) objectives. In addition, 11 Likert-type items assess student attitudes toward communication (i.e., willingness and confidence; ATC; Ball et al, 2014).

- Objective 1: Explain the fundamental processes that significantly influence communication.
- Objective 2: Construct messages consistent with the diversity of communication purpose, audience, context, and ethics.
- Objective 3: Respond to messages consistent with the diversity of communication purpose, audience, context, and ethics.
- Objective 4: Utilize digital literacy skills expected of ethical communicators.

3.2. Item type selection

The TOCS-2 consists of selected-response format items, with the number of response options ranging from four to five. Many items include video or audio components as part of the item stem. Items were written to motivate and engage test takers while easing resource commitments through electronic scoring. The distribution of media items among the learning objective was intended to reflect the activity level of the skill being measured. The affective, Likert-type (5-point scale) ATC items are included for purposes of assessing students' communication willingness and confidence.

3.3. Item pool and scale development process

In 2004, a group of Communication Studies faculty at James Madison University created the first Test of Oral Communication Skills (TOCS-1) based on core communication standards defined by the National Communication Association (NCA, 1998). At that time, technology to use audio-video and audio-only scenarios within question stems was just emerging, providing the opportunity to include innovative items in the original test. Moreover, given that video clips can easily portray actual communication behaviors, the use of video and audio is particularly relevant to tests of oral communication competence.

Development of the TOCS-2 had three goals: 1) update the video clips from TOCS-1; 2) replace or revise underperforming items; and 3) create more innovative items utilizing a combination of new user interfaces. Faculty first identified confusing or outdated items. The subject matter experts also determined that, although NCA has offered no revisions to the basic oral communication competencies since 1992, TOCS-1 underrepresented interpersonal communication, listening and information literacy as they relate to oral communication. Hence, revised student learning objectives were written (See Table 1). TOCS-2 now reflects the fundamental aspects of communication taught in most general education programs and found in most popular texts.

Table 1

Test Blueprint for TOCS-2 The numbers, below, reflect the most recent 100 items

TOCS-2 Objectives	# Items	Items
Objective 1: Explain the fundamental processes that significantly influence communication	25	2, 3, 4, 5, 6, 9, 11, 12, 14, 19, 24, 42, 61, 62, 63, 64, 67, 68, 69, 70, 71, 72, 73, 74, 75
Objective 2: Construct messages consistent with the diversity of the communication purpose, audience, and context.	25	7, 8, 10, 13, 21, 25, 36, 37, 38, 39, 40, 46, 47, 48, 49, 50, 51, 52, 53, 55, 56, 57, 65, 66, 89
Objective 3: Respond to messages consistent with the diversity of the communication purpose, audience, and context and ethics.	25	1, 15, 16, 17, 18, 20, 26, 27, 29, 30, 31, 32, 33, 34, 35, 44, 45, 58, 59, 60, 86, 87, 88, 99, 100
Objective 4: Utilize information literacy skills expected of ethical communicators.	25	22, 23, 28, 41, 43, 54, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98
Total Test	100	
Attitudes Toward Communication	# Items	Items
Willingness	6	1, 2, 3*, 4, 5, 6
Confidence	5	7, 8*, 9, 10, 11
Total	11	

* indicates reverse scored items.

A team of communications subject matter experts who teach in the university's general education program developed the final set of TOCS-2 items to align with the objectives in Table 1. The experts worked toward the development of items that were free of domain-specific content knowledge and that would engage students in application of the knowledge, skills and abilities gained through their communications course. For example, one of the media items asks the students to evaluate whether or not the speaker in the video cited primary or secondary sources. Throughout the process, the team of experts iteratively evaluated item content in tandem with empirical item analyses. Following individual subject matter expert's backward mapping to the standards, the group of subject matter experts carefully mapped the items to NCA standards via consensus.

Items on the Attitudes toward Communication scale (ATC) were developed for purposes of assessing affective components of communication. Specifically, ATC items align with two specific themes: willingness to communicate and confidence to communicate. Willingness items measure examinees' openness to communication, whereas confidence items address self-efficacy in communication. A confirmatory factor analysis of scores from the 11-item scale provided support for two related, yet distinct factors (Ball et al., 2014).

3.4. Subscores and their development

The test blueprint for the TOCS-2 and ATC items appears in Table 1. The first administration of the TOCS-2 was performed in Fall 2013. Internal consistency reliability for the overall test was reasonable (stratified coefficient alpha = .752). Use of subscale scores is not recommended at this time, due to lower than desirable levels of internal consistency (i.e., less than .50). Section Five contains a summary of Technical Information related to the Fall 2013 ($n = 2157$) administration of the test to first-year undergraduate students.

The ATC consists of two subscales; therefore, total scores should not be computed. Rather, individual subscale scores (i.e., willingness and confidence) should be computed by either summing or averaging student responses to the respective subscale items. Reliability estimates computed from general education students' scores were .77 (pre-test) and .83 (post-test) for the willingness subscale, and .72 (pre-test) and .71 (post-test) for the confidence subscale.

Section 4. Administrative Procedures

4.1. Proctor qualifications and training

Although administration of the TOCS-2 does not require intense training, proctors should be given guidance on standardized test administration as well as contemporary computer actions such as video playback and troubleshooting. Proctors should read this test manual prior to their training, which can be accomplished in a brief session covering test instructions and general test administration procedures. During training, proctors should carefully review the standardized instructions to be used in the actual testing session.

4.2. Testing procedures

TOCS-2 must be administered in a computer-based format. Examinees should be provided headphones and seated in a way that reduces distraction due to the visual media on other examinees' computer screens. Before beginning the test, examinees should be provided with general information about the number, type and content of items on the test. Examinees should be informed of the amount of time they will be given to complete the test and what they should do upon completion of the test. It is recommended that students be given at least 100 minutes to complete the TOCS-2. However, if the testing time is 100 minutes and the majority of students are still working after 90 minutes, the proctor may decide to extend the testing period for another 5 minutes. When the testing time is almost over (for example, at 95 minutes), the proctor should periodically announce the time remaining (e.g. 5 minutes, 2 minutes, 1 minute). ATC items are administered after all cognitive items have been completed.

The cognitive items for the TOCS-2 are randomly selected from an item bank. Random administration of items is for two main reasons: 1) test security, and 2) to prevent simultaneous accessing of the media items across many test takers (i.e., minimizing band-width strain). Items were numbered as they were developed. As a result, the numbering indicated in Table 1 is not reflective of the item's delivery on the test. Rather, the number refers to the unique ID of that item in the bank.

4.2.1. Computer-based administration

Students will each be seated in front of a computer on which the TOCS-2 test has been loaded. Students should receive either oral or written instructions on how to enter their identification information and begin the test. These standardized instructions are read immediately preceding administration of the TOCS-2; all data generated for this manual were collected in conditions which utilized these instructions. The instructions which follow should also be provided on the computer screen just prior to administering the first item.

“You are about to take a test to demonstrate your oral communication proficiency. You will be given 100 minutes to respond to 100 multiple choice questions and complete an 11-item attitude survey. If you have a question during the testing period, raise your hand and a proctor will come to you. An announcement will be made when there are 5 minutes remaining. When you have finished the test, please turn off your monitor, and inform the proctor that you are finished; you may leave once the proctor has dismissed you.”

Section 5. Technical Information

5.1. Scoring and interpretation

All TOCS-2 items are selected-response. When creating a total score for an examinee, correct responses to items are given a score of '1' and incorrect responses to items are given a score of '0.' The correct responses are then summed. The possible range of scores is therefore 0 to 100. A high total score would suggest that an examinee has high levels of knowledge about oral communication skills, and a low total score would suggest that an examinee has low levels of knowledge about oral communication skills. The Attitudes Toward Communication (ATC) items may be averaged by subscale (i.e., willingness and confidence). The range of possible averaged scores is 1 to 5 for both subscales. High ATC subscale scores suggest high levels of the intended construct.

5.1.1. Standard Setting

During the summer of 2014, 18 faculty members were recruited from a mid-Atlantic university to conduct standard setting procedures. The purpose of the establishment of a standard was to better inform test score interpretation and to arrive at a performance level faculty would consider acceptable. The faculty were selected to represent a cross-campus faculty cohort; thus, the faculty included three reference librarians and 15 faculty from across the university (e.g., foreign languages, writing, human communication, art history, business).

Standard scores were created that ranged from 200 to 800. A variant of the Mapmark standard setting procedure (Schulz & Mitzel, 2011) was used. The Mapmark procedure follows the earlier developed Bookmark procedure (Mitzel, Lewis, Patz, & Green, 2001) in Round 1 of the process but has been modified for Rounds 2 and 3. After all three rounds of faculty judgments were completed, the final standard was agreed to be a standard score of 480. This translates to a raw score of 78 correct, or 78% correct.

5.2. Evidence of reliability

Reliability refers to the degree of stability and consistency of test scores. Due to the numerous sources of variability in test scores, there are different ways of measuring reliability. The TOCS-2 has been examined for reliability as measured by Cronbach's alpha (α), which is frequently used to determine internal consistency. Specifically, α requires only one administration and is the mathematical equivalent of the average of all possible split-half reliability computations. Alpha indicates how much variance in the observed scores is attributable to the true score. Although coefficients with a value of .70 or higher have traditionally been considered adequate for scale use, reliabilities above .80 are desirable (Nunnally, 1978). The reliability (expressed using Cronbach's α) of the overall test scores from the Fall 2013 administration of the test is presented in Table 2; data collected in Spring 2014 paralleled the general distribution, mean scores and reliabilities as Fall 2013. Cronbach's alpha reliabilities were adequate for the total test; however, reliability estimates for the separate objectives were lower than desired, hence the caution in section 3.4 regarding the use of TOCS-2 subscale scores.

Table 2

Fall 2013 Cronbach's Alpha, Raw Scores, and Standard Deviations for TOCS-2

	<i>N</i>	α	<i>M</i>	<i>SD</i>
Total Test	2157	.75	79.92	7.27
<i>Analyses by Objective</i>				
Objective 1	2157	.57	81.55	10.88
Objective 2	2157	.45	80.63	9.77

Objective 3	2157	.48	77.07	10.89
Objective 4	2157	.27	74.45	7.33

Spring 2014 Cronbach's Alpha, Item-level Means, Standard Deviations for the ATC Subscale Scores Before and After Completing a General Education Human Communications Course

Subscale	Administration	N	α	M	SD
Willingness	Before Course		.77	4.19	.47
	After Course	70	.83	4.24	.60
Confidence	Before Course		.72	3.37	.66
	After Course	70	.71	3.69	.63

5.3. Evidence of validity

Validity refers to the degree to which one can make inferences from the scores obtained on a test. Validity is not an absolute state, but rather a collection of evidence indicating that the inferences based on the scores obtained on a test are valid for their intended use (AERA, 2000). As use of the TOCS-2 continues, the aim is to collect a body of evidence that the test items represent proficiency in human communication skills and that the scores distinguish students with knowledge/competency in introductory communication from those who do not.

Faculty who teach in the communication studies department wrote the TOCS-2 items using the objectives in conjunction with the National Communication Association (NCA) standards as their guide. All items on the TOCS-2 are aligned with the learning objectives outlined in Table 1.

Given that the TOCS-2 is a measure of achievement of general education in oral communication, it is realistic to expect students who earn higher scores in the SAT-Verbal test and higher courses grades to also perform better on the TOCS-2 than students with lower grades. That is, the TOCS-2 scores should be positively correlated with both the SAT-Verbal scores and grades in general education communication courses. One instructor made grades available for a final speech assignment, which was positively and moderately related to the TOCS-2 total scores (See Table 3). SAT-verbal scores were both statistically and practically related to TOCS-2 test scores, explaining 29.6% of the variance in test scores. Scores from the culminating course assignment (a performance assessment) was positively-related to TOCS-2 total scores.

Table 3

Correlation of TOCS-2 Scores with Selected External Variables (Fall 2013)

External Variable	r	p	r ²
SAT-Verbal (n = 1884)	.544	< .001	.296
Final Speech Assignment (n = 45)	.232	>.05	.054

Users of the TOCS-2 may want to consider the appropriateness of combining data from special populations, such as international students, with traditional domestic undergraduate students. At the institution in which the test was developed, international students enrolled in a pathways course (average 63.8% correct) scored significantly lower on the TOCS-2 than directly enrolled undergraduate students (average 80.8% correct), $t(2028) = 8.28, p < .001, d = 2.38$. One explanation for the difference is that the items contain culturally-specific scenarios, which international students may find more difficult than traditional domestic students. At the institution in which the test was developed, data from international students enrolled in a pathways program are analyzed separately from directly-enrolled students, and the findings are used to enhance instruction by focusing on specific areas on which students had difficulty on the test.

The Attitudes toward Communication (ATC) scores were written to address students' willingness and confidence in their communication skills. Ball and colleagues (2014) examined the factor structure of scores from the ATC and found support for the two factors. A sample of 581 community college students completed an early version of the TOCS-2 test and the ATC scale. Table 4 summarizes the relationship between the two ATC subscales (willingness and confidence) and the TOCS-2 for the community college sample. Note that the

two ATC subscales are positively related, yet distinct ($r = .445$). The ATC scores were also moderately, positively, and significantly related to the TOCS-2 (early version) scores. Further, note in Table 2 that average willingness scores did not change from before to after the course, yet average confidence scores increased by over a point.

Table 4

Correlation between ATC subscales and early version of TOCS-2 (Community College Sample)

Variable	Willingness	Confidence	TOCS-2
Willingness (ATC)	1.00		
Confidence (ATC)	.445**	1.00	
TOCS-2 (earlier version)	.377**	.245**	1.00

$N = 581$, ** $p < .01$

5.4. Norming

At the time of this writing, the TOCS-2 had been administered as a final examination to first-year students enrolled in the Fall 2013 general education human communication course ($n = 2157$). Gender of the students in the sample roughly approximated those of the overall undergraduate student population at the institution of study (61% female and 34% male, 5% unknown).

To determine how the students at your institution performed in relation to college students at the institution that serves as the site of TOCS-2 research, refer to Table 5, which contains percentile rank information for 2013-2014 first-year students. As an example of how to interpret a percentile rank, suppose that a student achieved a scores of 92 on the TOCS-2. Based on this user group data, 98.19% of students scored at or below a 92, and a student with a score of 92 would be at the 98th percentile.

Table 5

*Percentile Ranks for TOCS-2 Scores for **First-Year Students** at a Mid-Atlantic 4-Year Institution*

Score	Total Group ($N = 2157^*$)
96	100
95	99.77
94	99.44
93	99.07
92	98.19
91	96.94
90	95.36
89	93.14
88	90.54
87	87.20
86	82.34
85	76.82
84	72.09
83	66.25
82	60.32
81	54.38
80	48.59
79	43.16
78	37.51
77	32.73
76	28.47
75	24.99

74	21.46
73	18.31
72	15.11
71	12.52
70	10.06
69	8.44
68	6.77
67	5.38
66	4.36
65	3.48
64	2.69
63	2.41
62	1.90
61	1.48
60 or fewer	1.11

Section 6. Additional Information

6.1. Future work to be conducted

To enhance test use and interpretation for future administrations, the following studies or analyses are recommended:

- Evaluation of instrument with populations other than first-year students at a mid-Atlantic 4-year institution, including differential item functioning for special populations (e.g., international students);
- Conduct an Exploratory Factor Analysis (EFA) to investigate the structure of the TOCS-2;
- Use an iterative process to increase reliability. In particular, consider:
 - Additional administrations
 - Item analyses
 - Content revision

6.2. Where to find additional information

At the time of this writing, the TOCS-2 had been administered to one sample of students ($n = 2,157$) at James Madison University. Additional information on the TOCS-2 may be obtained by contacting Dr. Timothy Ball, Assistant Professor of Communication Studies, James Madison University, (balltc@jmu.edu) or the Center for Assessment and Research Studies, James Madison University, MSC 6806, Harrisonburg, VA 22807.

Section 7. References

- American Educational Research Association, American Psychological Association, & National Council of Measurement in Education. (2000). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Ball, T. C., Jurich, D. P., & Anderson, R. D. (2014). *Developing a measure of students' attitudes towards communication*. Manuscript submitted for publication.
- Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Standard setting: Concepts, methods, and perspectives* (pp. 249–281). Mahwah, NJ: Erlbaum.
- National Communication Association (1998). *Speaking and listening competencies for college students*. Washington, DC: National Communication Association.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Schulz, E. M. & Mitzel, H. C. (2011). Understanding Rasch measurement: A Mapmark method of standard setting as implemented for the National Assessment Governing Board. *Journal of Applied Measurement*, 12, 165-193.