

# **The Information Literacy Test (ILT)**

## **Test Manual**



**Madison Assessment**

[www.madisonassessment.com](http://www.madisonassessment.com)



THE CENTER FOR  
**ASSESSMENT &  
RESEARCH**  
STUDIES

Revised April, 2014  
By Matthew Swain, Donna L.  
Sundre & Kathy Clarke

MSC 6806  
Harrisonburg, VA 22807  
540.568.6706 Phone  
540.568.7878 Fax  
[assessment@jmu.edu](mailto:assessment@jmu.edu)  
[www.jmu.edu/assessment](http://www.jmu.edu/assessment)

---

**Table of Contents**

## **The Information Literacy Test (LMT)**

<b>Section 1. The Nature of the Instrument</b>	<b>3</b>
<b>Section 2. Intended Use</b>	<b>3</b>
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
<b>Section 3. Test Development</b>	<b>3</b>
3.1 <i>Academic and theoretical basis</i>	3
3.2 <i>Item type selection</i>	4
3.3 <i>Item pool and scale development process described</i>	4
<b>Section 4. Administrative Procedures</b>	<b>7</b>
4.1 <i>Proctor qualifications and training</i>	7
4.2 <i>Testing procedures</i>	7
4.2.1. Security	7
4.2.2. Process of Administration	7
<b>Section 5. Technical Information</b>	<b>8</b>
5.1 <i>Scoring and interpretation</i>	8
5.2 <i>Evidence of reliability</i>	8
5.3 <i>Evidence of validity</i>	9
5.4 <i>User-Group Data</i>	10
5.5 <i>Proficiency Level Standard Setting</i>	12
<b>Section 6. Additional Information</b>	<b>13</b>
6.1 <i>Where to get additional information</i>	13
6.2 <i>Work to be conducted to enhance test use and interpretation</i>	13
<b>Section 7. References</b>	<b>13</b>
<b>Section 8. Appendix</b>	<b>13</b>
<i>Notes for proctors</i>	13

## The ILT Manual

### Section 1. The Nature of the Instrument

The Information Literacy Test (ILT) is a 60-item multiple-choice test developed by librarians and assessment specialists. The ILT is based on the Association of College & Research Libraries (ACRL) Information Literacy Competency Standards (See [www.ala.org/acrl/standards/informationliteracy](http://www.ala.org/acrl/standards/informationliteracy)). This instrument was designed to directly assess collegiate students' competencies in information literacy.

### Section 2. Intended Use

#### 2.1 Appropriate and inappropriate uses and interpretations

This instrument was designed to assist institutions in identifying students' abilities to "locate, evaluate, and effectively use information when it is required" (ACRL, 2003). It measures the information literacy skill of anyone who should have such skills including college and graduate school students. The ILT was developed for use at the programmatic level. Thus, inferences made about learning or mastery should be made *only* in the aggregate.

The ILT was *not* designed for making decisions about individual students. Currently, its psychometric properties are not sufficient to support high-stakes classifications for individuals (please refer to section 5.2 -- Evidence of reliability). This instrument was also not intended as a vehicle for providing individual students with feedback about their mastery of information literacy skills. Institutions may choose to provide their students with individual feedback, but results should *not* be used to make high-stakes classification decisions. According to the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2000), test users are responsible for collecting validity evidence for any uses of the test other than those recommended here. In addition, other institutions are encouraged to explore score reliability and to validate the inferences they wish to make at their home institution.

The data collected with the ILT can be used to provide information about student learning that can inform improvements to information literacy programming. The results of the ILT can and have been successfully used to meet accountability demands. For example, the State Council of Higher Education in Virginia's (SCHEV) has mandated that all public institutions must report on student competencies in the area of technology/information literacy. The information submitted to SCHEV received high praise. The ILT has enjoyed widespread use at over 40 institutions around the globe.

#### 2.2 Target population

The primary focus during test development was on college students, whether graduate or undergraduate, enrolled at either a four-year university or community college. To determine if the ILT is appropriate for a particular population, one should consider the learning objectives the ILT was created to assess. See Table 1 for a list of those standards. If these standards sufficiently align with the learning objectives of the information literacy program in question, the test may be considered appropriate. In addition, a potential test user should consider examining the items. Again, if the items appear to be covering the appropriate topic areas and skills, and if they appear to be written at an appropriate level of difficulty, the ILT could be considered appropriate for that particular population and use.

#### 2.3 Qualifications of users

Test users must be trained to administer assessments in standardized conditions. The Proctor qualifications and training section of this manual (Section 4.1) provides more information about how proctors can be trained for test administration. In addition, test users should be knowledgeable about how to interpret the statistical results from the test and how to make appropriate inferences about the program using the results. Test users who do not have a measurement background or do not have in depth knowledge of the program are encouraged to consult with colleagues who have the necessary knowledge.

### Section 3. Test Development

#### 3.1 Academic and theoretical basis

The ILT was designed to evaluate student learning in four content areas of the ACRL Information Literacy Competency Standards for Higher Education. The first has to do with defining and articulating the nature and extent of information needed. The second standard focuses on whether students can efficiently access and use needed information. The third

objective speaks to students' ability to evaluate information and its sources critically and incorporate selected information into his or her knowledge base and value system. The fourth objective specifies evaluation of students' ability to use information to accomplish a specific purpose. The last standard focuses on the student's understanding of the ethical, legal, social, and economic issues surrounding the use of information and information technology.

### 3.2 Item type selection

All ILT items are selected-response. The items were written as such to ease scoring, to maintain objective scoring, and to minimize test-taker fatigue. Most items follow a typical multiple choice format, in which an item stem is followed by alternative responses consisting of the correct answer and several distracters. The number of alternative responses to each item on the ILT range from three to six.

### 3.3 Item pool and scale development process described

There were several criteria that guided the development of the ILT. It was to be a multiple-choice test that should be completed within one hour; however we do not recommend a rigid time limitation. The items were to make frequent use of graphics and it was anticipated that approximately two thirds of the items would measure lower-order skills with the remaining one third measuring higher-order skills (as defined in the Information Literacy Competency Standards). In addition, The ILT was to be web-based (i.e., administered over the Internet).

The ILT was originally intended to measure all five ACRL standards. However, Standard Four was inappropriate for a multiple-choice item format. This standard refers to the student's ability to use information effectively to accomplish a specific purpose. This standard would be more reasonably assessed through review of student developed products, processes and performances. Because of this constraint, the ILT items were developed to measure Standards One, Two, Three, and Five. Our librarians also determined that Standards Two and Three should receive greater emphasis on the test. Consequently, these two standards comprise approximately two thirds of the test, and Standards One and Four contributed to the remaining third in roughly equal proportions.

During the first half of 2003, several university reference librarians developed and revised an initial pool of 80 items. The numbers of response options for these items ranged from two to five, with most of the items having four or five options. These items comprised the pilot form of the ILT, which was then entered into the Adaptex test administration software (Wise & Yang, 2003) and administered to a sample of 506 incoming first-year students at a moderate-sized southeastern public university. Based on an item analysis of the data from the pilot form, 60 items that exhibited good psychometric properties were selected for use on the final ILT form. The 60-item ILT was subsequently administered to a random sample of 524 mid-year sophomores in February, 2004 during the university's Assessment Day. Over time, all items have received additional review; for example in 2013, our librarians slightly modified 14 items to enhance their clarity and validity.

The final version of the ILT contains 60 multiple-choice items. The final test blueprint outlining the specifications for the ILT is shown in Table 1. A more detailed item mapping for each standard is presented in Tables 2 through 5 below.

Table 1. Test Blueprint for ILT

Scales	# of Items	Items as numbered on the ILT
<u>Standard 1</u> : defines and articulates the nature and extent of information needed.	8 13% of test	1,4,5,7,10,11,43,57
<u>Standard 2</u> : accesses needed information effectively and efficiently	24 40% of test	2,3,8,9,13,14,15,16,17,18,20,21,22,23,24, 25,26,27,28,29,31,33,34,35
<u>Standard 3</u> : evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system.	19 32% of test	6,12,19,30,32,36,37,38,39,40,41,42,44,45 46,47,48,49,50
<u>Standard 5</u> : understands many of the ethical, legal, and socio-economic issues surrounding information and information technology.	9 15% of test	51,52,53,54,55,56,58,59,60
Total Test	60 100% of test	1-60

Table 2. Individual Standards for ILT

ILT	General Item Level Description	Overall Description of Items
<p><b>Standard 1:</b> defines and articulates the nature and extent of information needed.</p>	<ul style="list-style-type: none"> <li>• 1. Knowing the appropriate resource for a topic</li> <li>• 2. Knowing effective search terms</li> <li>• 4. Knowing the appropriate resource based on depth of information needed</li> <li>• 5. Distinguishing between types of sources</li> <li>• 6. Distinguishing between types of sources</li> <li>• 7. Distinguishing between types of sources</li> <li>• 8. Knowing where sources are located</li> <li>• 9. Acquiring a source</li> <li>• 10. Narrowing information</li> <li>• 11. Distinguishing between sources</li> <li>• 12. Knowing what “peer review” is</li> <li>• 49. Distinguishing between references (aka works cited)</li> </ul>	<p>Items mapped to standard one broadly assess students’ knowledge of identifying and distinguishing between different types of sources (i.e., research article, encyclopedia, and dictionary). This includes the ability to identify the correct source for retrieving needed information, knowing how to use efficient search keywords, knowing how to locate physical sources (i.e., microform, periodicals, circulations), the knowledge to narrow down information, and the ability to identify the type of source a reference belongs to (e.g., journal article vs. book).</p>

Table 3. Individual Standards for ILT

ILT	General Item Level Description	Overall Description of Items
<p><b>Standard 2:</b> accesses needed information effectively and efficiently</p>	<ul style="list-style-type: none"> <li>• 13. Database querying</li> <li>• 14. Distinguishing between databases</li> <li>• 15. Searching for publications</li> <li>• 16. Searching for publications</li> <li>• 17. Knowledge of search operators</li> <li>• 18. Knowledge of finding books in a library</li> <li>• 19. Knowledge of citing</li> <li>• 20. Knowledge of search operators</li> <li>• 21. Knowledge of search operators</li> <li>• 22. Knowledge of search operators</li> <li>• 23. Accessing a publication</li> <li>• 24. Database querying</li> <li>• 25. Knowledge of reference types</li> <li>• 26. Knowledge of reference types</li> <li>• 27. Knowledge of reference types</li> <li>• 28. Knowledge of reference types</li> <li>• 29. Knowledge of reference types</li> <li>• 30. Knowledge of reference types</li> <li>• 31. Knowledge of the bibliography references (aka works cited)</li> </ul>	<p>Items mapped to standard two broadly assess students’ knowledge and ability to access and use information from search results in an index database. This includes knowledge of querying and using keywords, knowledge of the search process, knowledge and use of search operators (e.g., AND, OR, wild cards), knowledge of proper citation, and the ability to distinguish between different types of publications based on a provided reference (e.g., journal article vs. book).</p>

Table 4. Individual Standards for ILT

ILT	General Item Level Description	Overall Description of Items
<p><b>Standard 3:</b> evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system.</p>	<ul style="list-style-type: none"> <li>• 3. Refining the search / Knowledge of search operators</li> <li>• 32. Evaluating source credibility</li> <li>• 33. Using data from a table</li> <li>• 34. Using data from a table</li> <li>• 35. Using data from a table</li> <li>• 36. Using data from a table</li> <li>• 37. Evaluating a claim</li> <li>• 38. Using information</li> <li>• 39. Evaluating source credibility</li> <li>• 40. Knowing a sources purpose (e.g., stating facts, persuasion, etc.)</li> <li>• 41. Identifying sources author</li> <li>• 42. Evaluating source credibility</li> <li>• 43. Evaluating source credibility</li> <li>• 44. Evaluating a claim</li> <li>• 45. Identifying sources author</li> <li>• 46. Knowledge of source types</li> <li>• 47. Using information</li> <li>• 48. Selecting an appropriate source</li> </ul>	<p>Items mapped to standard three broadly assess students' ability to use and critically evaluate source information. This include evaluating the credibility and reliability of a source, extracting information from data presented in a table, evaluating a source's claims, awareness of the purpose of a source (e.g., persuasion vs. factual), the ability to identify the author a source, the ability to draw the appropriate conclusion from information provided from a source, the ability to identify the type of source that will best answer a provided question.</p>

Table 5. Individual Standards for ILT

ILT	General Item Level Description	Overall Description of Items
<p><b>Standard 5:</b> understands many of the ethical, legal, and socio-economic issues surrounding information and information technology.</p>	<ul style="list-style-type: none"> <li>• 51. Knowledge related to source availability and/or access</li> <li>• 52. Knowledge related to source availability and/or access</li> <li>• 53. Knowledge of ethical/legal issues of using sources</li> <li>• 54. Knowledge of ethical/legal issues of using sources</li> <li>• 55. Knowledge of ethical/legal issues of using sources</li> <li>• 56. Knowledge of ethical/legal issues of using sources tables/graphs</li> <li>• 57. Knowledge of resources for proper citation</li> <li>• 58. Knowledge of creating a reference</li> <li>• 59. Knowledge of ethical/legal issues of sharing an audio source</li> <li>• 60. Knowledge of citing direct quotes</li> </ul>	<p>Items mapped to standard five broadly assess student's understanding of ethical and legal issues regarding the use of information from a source. This includes knowledge of the availability/accessibility of different types of sources, knowledge of information that is freely available vs. information that is proprietary, knowledge of ethical and legal issues regarding the use of information extracted from a source, knowledge of creating appropriate references and citations, and knowledge of sharing intellectual property.</p>

## Section 4. Administrative Procedures

### 4.1 Proctor qualifications and training

While administration of the ILT does not require intense training, proctors should be given guidance on standardized test administration. Proctor training can be accomplished in a brief session in which they are familiarized with the test instructions and the procedures to follow during test administration. During training, proctors should be provided with the standardized instructions to be used in the actual testing session. General instructions for administration are provided in the following section.

### 4.2 Testing procedures

The ILT is currently administered online via a URL provided by Madison Assessment. Detailed instructions are provided with the URL and password when the login document is sent.

#### 4.2.1. Security

Because the ILT is currently being administered in many different settings, security is of utmost importance. Specifically, caution has been taken to ensure that these items are secure and we ask that test users take the same caution when they have access to the items. The following guidelines will ensure the security of the testing program and alleviate test and score validity issues. These guidelines are also included in the Ordering Agreement institutions sign prior to test use.

- All tests will be administered in a proctored test environment.
- Students, proctors, or other individuals will not leave the testing session with any recorded information about the content of the test. This includes scratch paper or notes taken during testing, and the ILT web address and passwords. Cell phones cannot be used in any way during testing.
- The web address and password for the ILT will not be shared with anyone other than those who need to have access to the test (i.e., examinees and proctors).

Depending on the situation in which the test is administered, other security concerns may arise. If there are questions regarding how to handle security in particular testing situations, please contact Madison Assessment. In addition, please report any situations, which indicate a problem with the security of the test itself.

#### 4.2.2. Process of Administration

The ILT is offered as a web based test. A unique URL is provided to each client. Test-takers enter a secure website and provide identifying information and a password to enter the secure software. At the beginning of the test the following testing instructions are presented to students.

The ILT in its current version consists of 60 scored items. Traditionally, most students are able to complete all 60 items within a 60 minute time frame. However, the testing time allotted for a particular administration can be determined by the administrator. The ILT is not meant to be a speeded test (where students work quickly to see how much they can complete in a given time); however, it is important that students realize that there is a time limit so they will stay focused and on task.

Each item is presented on a separate page. Each window presents a single item's stem and response options. When students are ready to respond, they should click the circular button next to the desired response option and then click the *Next* button. Before moving to the next item, a response must be given for the current item. After the examinee has progressed through all 60 items, they will be asked whether they would like to review their responses prior to final submission.

The examinee's home institution also has several options that might be provided to students upon completion of the test. They might be presented with their score on the test (in terms of percentage correct). The testing software can be programmed to show or suppress the test taker's individual score. They may also print out a certificate of test completion to present to their class instructor. The testing institution will be provided with a data file containing students scored responses and total scores for each examinee. This data file is provided in an EXCEL format. The scored data files for each institution will be sent when the agreed upon test administration window has passed.

## Section 5. Technical Information

### 5.1 Scoring and interpretation

All ILT items are selected response. The majority of items have three response options including the correct response. The range is between three and six response options. Three response options are considered the optimal number of choices for multiple-choice test items (Rodriguez, 2005). Items are scored dichotomously: a correct response to an item is given a score of '1' and an incorrect response to an item is given a score of '0.' The total score is obtained by summing the scored item responses. Higher total scores indicate that examinees have higher levels of information literacy, and lower total scores indicate that examinees have lower levels of information literacy.

### 5.2 Evidence of reliability

An important feature of any psychometric test is that it produces test scores that show sufficient reliability. The 60-item ILT was administered to a random sample of 1,035 students from 2013-2014 at various 2- and 4-year institutions. Table 6 shows the means, standard deviations, and internal consistency estimates (coefficient alphas) for the total ILT as well as for the items from each of the standards. On average, all students passed nearly 64% of the items, and exhibited a strong score internal consistency for the total score. The internal consistency estimates for the standards (subscales), while not as high, are adequate given the parameters for each standard. As noted in the table, several of these standards are represented by a small number of items which can result in lower estimates of reliability. Because the reliability of these subscales are somewhat lower it is recommended that scores associated with the subscales be interpreted **only** at a group or descriptive level; in other words, these scores should not be used to make interpretations about the unique strengths or weaknesses of an individual student.

Table 6. Descriptive statistics and reliability estimates for the 2013-2014 ILT administrations

Scale	# Items	2-Year		4-Year	
		Mean (SD)	Coefficient Alpha	Mean (SD)	Coefficient Alpha
Total Test	60	38.39 (8.57)	0.87	36.68 (10.71)	0.91
<u>Standard 1</u> : defines and articulates the nature and extent of information needed.	8	5.38 (1.47)	0.45	5.41 (1.75)	0.61
<u>Standard 2</u> : accesses needed information effectively and efficiently.	24	12.92 (3.58)	0.68	12.66 (3.83)	0.71
<u>Standard 3</u> : evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system.	19	13.90 (3.54)	0.77	12.84 (4.22)	0.84
<u>Standard 5</u> : understands many of the ethical, legal, and socio-economic issues surrounding information and information technology.	9	6.19 (1.68)	0.45	5.77 (2.19)	0.67

Two-Year: N = 945; Four-Year: N = 74

Though the data trends provide further reliability and generalizability evidence for the use of the ILT, there are limitations to the inferences that can be made from these results. Comparability of institutions within the four-year and two-year institutions and comparability between both types of institutions is limited due to the lack of standardized data collection in methods and testing environments. Mean scores, standard deviations, and reliability coefficients are sample dependent. Without further detail about test administrations and sampling techniques, comparisons should be made with caution.

For example, review of Table 6 indicates that results for the 2-year institutions are slightly higher than those for the 4-year institutions. This may be due to the fact that many 2-year institutions administer the ILT only to students who are graduating with their Associate in Arts (A.A.) or Sciences (A.S.) degrees, while the 4-year institutions frequently employ a pre and post-test design. Without more information concerning how subjects were sampled, further interpretation of results is cautioned.

### 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 scores obtained on a test are valid for their intended use (AERA, 2000). For the ILT, two types of evidence have been collected: that based on expert ratings of the items (content validity) and that based on the degree to which ILT scores statistically behave as we would expect a measure of information literacy to behave (construct validity).

To assess content validity, three university reference librarians were provided descriptions of the four ACRL standards measured by the ILT as well as the standard each item was intended to measure. The librarians then studied each ILT item and independently rated the extent to which the item matched its purported standard using three rating categories: “Matched the Standard,” “Uncertain,” or “Did Not Match the Standard.” The item alignment rating results were highly favorable, as all three librarians agreed on 42 of the 60 items (70%), and at least two agreed on 59 items (98%). These results indicate that ILT items displayed content validity through alignment to the intended ACRL standards.

Construct validity evidence was obtained through five additional studies. The first used the data from the initial administration of the ILT to university sophomores (spring, 2004 Assessment Day) described earlier. Most of those students had taken the Information Seeking Skills Test (ISST), another Information Literacy Test, as first year students. The ISST was developed as a high stakes competency test for use only at James Madison University. The instrument measures skills and reference materials held at JMU. ISST scores were obtained for 333 students by Miller (2004) who found the correlation between the ILT and the ISST to be positive and significant [ $r(331) = .38, p < .001; r^2 = .144$ ]. A problem with data collected on Assessment Day, or other low-stakes testing conditions, is that some students may not try very hard because there are no personal consequences associated with test performance. Wise and Kong (2005) showed that item response times can provide a valid measure of the amount of effort a student devotes to a computer-based test. There were 36 students who did not exhibit effort on at least 95% of their items. These students were deleted from the sample, and the ILT-ISST correlation increased to .45 [ $r(295) = .45, p < .001; r^2 = .203$ ]. Both are reasonable effect sizes.

In the second study the ILT was administered to 121 introductory psychology students during the fall, 2004 semester, 75 of whom were first-year students and the remaining 46 were sophomores. Immediately after taking the ILT, students were administered an eight-item survey, which contained five questions regarding frequency of course-related information literacy activities and three questions regarding confidence in finding and evaluating information. The results showed that the sophomores scored significantly higher on the ILT than the first-years [ $t(119) = 2.06, p = .041, d = 0.39$ ]. In addition, ILT scores were significantly correlated with cumulative GPA [ $r(119) = .20, p = .032$ ]. The analysis of the survey items revealed significant correlations between ILT scores and two of the three confidence items: “confidence in ability to find books and scholarly articles for project of interest,” [ $r(119) = .33, p < .001$ ], and “confidence in ability to evaluate resources for their quality,” [ $r(119) = .26, p = .005$ ]. The ILT scores were uncorrelated with all of the items regarding frequency of course-related information literacy activities; however, similar findings for the ISST were reported by Cameron (2004). These results show sensitivity to college level experience.

The third study compared the ILT scores of 422 incoming first-year students—collected in fall, 2004—with the scores of 524 mid-year sophomores—collected in spring, 2004. The first-year group showed a mean of 37.13 and a standard deviation of 7.70, while the sophomore group showed a mean of 41.61 and a standard deviation of 8.45. The means were found to be significantly different [ $t(944) = 8.43, p < .001, d = 0.53$ ]. These results are consistent with the fact that the sophomores, unlike the first-years, had been exposed to instructional modules in information literacy and had demonstrated competency on the ISST. The  $d = .53$  indicates an effect size of more than one half standard deviation. This is an impressive effect size.

In a fourth study, year 2008 and year 2009 ILT data from a sample of 683 first-year students aggregated across four four-year institutions were compared with the ILT scores collected from the sample of 422 JMU first-years in 2004. The JMU group showed a mean of 37.13 and a standard deviation of 7.70, while the first-year students from the four four-year institutions showed a mean of 36.12 and a standard deviation of 7.71. The two groups were found to be significantly different [ $t(1103) = 2.11, p = .0035, d = 0.13$ ]. Though these groups differed significantly on their mean scores, it is important to note that the magnitude of that difference may be considered small. This magnitude is determined by Cohen’s  $d$ . Though Cohen provided rules of thumb for interpreting the value of  $d$ , he also noted the importance of the researcher’s judgment in determining the most appropriate interpretation (Cohen, 1988). These results indicate that entering first-year students from several different institutions were not dramatically different from JMU entering students. The significant differences observed were statistically different from zero, but the effect size of .13 is indicative of a slight difference.

ILT data collected in year 2008 and year 2009 from 839 first-year students aggregated across five two-year institutions were also compared with the year 2004 sample of 422 JMU first-years. The JMU group showed a mean of 37.13 and a standard

deviation of 7.70, while the first-year students from the five two-year institutions showed a mean of 35.77 and a standard deviation of 7.92. Again, the two groups were found to be significantly different [ $t(1259) = 2.90, p = .0037, d = 0.17$ ], but the magnitude of that difference again was relatively small. These results suggest that JMU entering students were not dramatically different from entering students from several two year institutions. Again, statistical significance was observed, but the effect size was again quite small.

These known group differences provide some evidence of construct validity. In each of the aforementioned studies, group differences were realized as expected. In the JMU samples, sophomore students who had been provided educational opportunities in information literacy scored higher than the first-years who had not experienced JMU's information literacy experiences. It was also expected that first-year students attending a more selective university, JMU, might perform better than first-year students at open enrollment institutions. Again, while the observed differences were statistically significant, it may be consoling to community colleges that their entering students did not perform remarkably different on information literacy upon entry. These results also strongly convey the importance of providing intentional instruction and assessment of information literacy. The clearly do not arrive with sufficient knowledge or competencies. *The primary issue institutional test users will be interested in is whether or not significant differences can be observed in their own students after experiencing relevant course work and experiences at their institutions. Results to date suggest that the ILT is sensitive to identifying these differences.*

Table 7 illustrates the data obtained from the four studies discussed above.

Table 7. Comparison of ILT mean scores across samples.

Sample	N	Mean	SD
JMU First-years 2004	422	37.13	7.70
JMU Sophomores 2004	524	41.61	8.45
Five Two-Year Institutions First-years 2008-2009	839	35.77	7.92
Four Four-Year Institutions First-years 2008-2009	683	36.12	7.71

In the fifth study, data was collected from a sample of 1,035 students from 2013-2014 at various 2- and 4-year institutions. Students were asked to indicate how many college credit hours they had completed. They were given eight credit hour categories to select from. The credit hour bands and associated descriptive statistics are provided in Table 8.

Table 8. ILT Total Scores Split by Credit Hour Band

Credit Hours	Freq	%	2-Year			
			Mean	SD	Min	Max
Less than 15	89	8.60%	36.56	9.34	13	53
16 to 30	114	11.01%	38.17	8.12	9	52
31 to 45	95	9.18%	38.62	9.86	11	53
46 to 60	187	18.07%	39.61	8.95	12	57
61 to 75	387	37.39%	37.99	8.24	11	57
76 to 90	64	6.18%	37.86	9.26	14	55
91 to 105	36	3.48%	36.69	9.18	14	49
106 or more	63	6.09%	38.67	8.49	16	56

N = 1,035

Collectively, the evidence obtained thus far supports the validity of ILT scores as measures of students' information literacy knowledge and skills. This conclusion is supported both by content- and construct-related validity findings.

## 5.4 User-Group Data

The following data was the same data collected in the Reliability and Validity sections. Students from various 2 and 4-year institutions completed the ILT between 2013 and 2014. To determine how the students at your institution performed in relation to students at an institution similar to yours, refer to Table 9.

Table 9. Percentile Ranks for ILT Total Scores by Institution Type

ILT Total Score	Whole Sample	2-Year	4-Year
9	0.10	0.11	--
10	--	--	--
11	0.29	0.32	--
12	0.39	0.42	--
13	0.48	--	1.35
14	0.87	0.74	2.70
15	1.45	1.16	5.41
16	1.74	1.48	--
17	2.42	2.22	--
18	2.80	2.54	6.76
19	3.48	2.96	10.81
20	4.25	3.70	12.16
21	4.93	4.23	13.51
22	5.70	4.76	17.57
23	6.57	5.71	--
24	7.73	6.98	--
25	9.18	8.57	--
26	10.92	10.26	18.92
27	11.50	10.79	20.27
28	13.91	13.44	--
29	16.33	15.56	25.68
30	18.74	18.10	27.03
31	21.26	20.63	28.38
32	23.38	22.86	29.73
33	26.86	26.35	32.43
34	30.14	29.42	36.49
35	34.01	33.44	--
36	37.58	37.14	37.84
37	41.16	40.74	41.89
38	45.51	44.97	45.95
39	50.53	49.95	52.70
40	55.36	54.81	56.76
41	60.29	59.68	62.16
42	65.12	64.66	66.22
43	69.76	69.31	71.62
44	74.11	73.76	75.68
45	79.52	79.26	81.08
46	83.57	83.39	85.14
47	87.34	87.30	87.84
48	90.24	90.37	89.19
49	92.66	92.80	91.89
50	94.40	94.71	--
51	95.75	96.19	--
52	97.49	97.46	97.30
53	98.55	98.52	98.65
54	99.32	99.37	--
55	99.61	99.68	--
56	99.81	99.79	100.00
57	100.00	100.00	--

Whole Sample: N = 1,034; Two-Year: N = 945; Four-Year: N = 74

These samples sizes are small (especially for four-year institutions) and should therefore be used with caution until more data are collected.

## 5.5 Proficiency Level Standard Setting

Without an interpretive context, test scores have little meaning. One way in which the ILT could be leveraged for greater benefit would be to administer the test to students at two different points in time and calculate the average change in scores. In this context, the posttest scores are interpreted relative to the pretest scores. This value-added approach to assessing student outcomes is often used in higher education assessment.

A limitation to the value-added approach, however, is that it does not provide information regarding the degree to which the students have learned as much as you expect or intend them to. An alternative to the value-added approach is to identify the absolute point on the test score scale corresponding to a particular level of proficiency. The most rigorous methods available for test developers to obtain this information are collectively termed *standard setting methods*. In standard setting, a panel of judges is provided a definition of one or more levels of proficiency and a copy of the test items. The judges are then asked to make judgments regarding the test scores that correspond to those levels of proficiency. This provides a more absolute context in which to interpret test performance.

A standard-setting workshop for the 60-item ILT was conducted during March 2004. A modified version of the Bookmark standard setting method (Lewis, Green, Mitzel, Baum, & Patz, 1998) was used that required two half-day sessions to complete. Ordered item booklets were compiled, using 43 ILT items whose Bookmark location values were computed from data from the fall, 2003 pilot testing of incoming JMU first-year students.

There was a diverse panel of 10 judges used in the workshop. Three were librarians from James Madison University, three were librarians from Virginia community colleges, one was a librarian at another Virginia university, two were faculty in our Center for Assessment and Research Studies (CARS), and one was a doctoral student in assessment and measurement.

Two performance standards were set. The first standard differentiated examinees who were considered *Proficient* from those that were *Below Proficient*. The second differentiated those who were *Advanced* from those that were *Proficient*. Prior to the workshop, definitions were created for what students should know and be able to do at the Proficient and Advanced levels. At the beginning of the workshop, participants discussed the definitions, which were then used by the judges as they made their judgments.

Table 10 shows the proficiency definitions given to the judges and the resultant performance standards the panel recommended. For the Proficient designation, the judges recommended a raw score performance standard of 39, which corresponded to 65% correct. For the Advanced designation the performance standard was 54, which corresponded to 90% correct. If these performance standards were applied to spring, 2004 administration of the ILT, the percentages of students in the Below Proficient, Proficient, and Advanced categories were 17, 77, and 4, respectively. Results like these can provide a clear interpretive benchmarks regarding how many students demonstrated adequate levels of proficiency in a particular sample. ILT test users are invited to either use these standards or to create their own community performance expectations.

Table 10. Performance level definitions and standards recommended for Proficient and Advanced levels on the 60-item ILT

Proficiency Level	Performance Standard	Descriptors
Proficient	39(65%)	<p><u>The student who is <i>Proficient</i> is able to:</u></p> <ul style="list-style-type: none"> <li>Describe how libraries are organized.</li> <li>Define major library services.</li> <li>Choose the appropriate type of reference source for a particular information need.</li> <li>Identify common types of citations.</li> <li>Employ basic database search strategies.</li> <li>Locate a variety of sources in a library or online.</li> <li>Discriminate between scholarly and popular publications.</li> <li>Legally and ethically use information.</li> </ul>
Advanced	54 (90%)	<p><u>The student who is <i>Advanced</i> is able to attain the criteria for Proficient and:</u></p> <ul style="list-style-type: none"> <li>Modify and improve database search strategies to retrieve better results.</li> <li>Employ sophisticated database search strategies.</li> <li>Interpret information in a variety of sources.</li> <li>Evaluate information in terms of purpose, authority and reliability.</li> <li>Understand ethical, legal, and socioeconomic issues relating to information access and use.</li> </ul>

It should be noted that these recommended performance standards are linked to the particular definitions we used in our standard setting workshop. They may provide meaningful interpretive benchmarks for other institutions that adopt our

performance definitions. For institutions adopting different definitions, however, a separate standard setting would be appropriate. Establishing community expectations for student performance provides a highly valued interpretive framework.

## Section 6. Additional Information

### 6.1 Where to get additional information

Additional information on the ILT may be obtained by contacting Madison Assessment ([info@madisonassessment.com](mailto:info@madisonassessment.com)). Information may also be obtained through the following website: [www.madisonassessment.com](http://www.madisonassessment.com).

### 6.2 Work to be conducted to enhance test use and interpretation

Future work with the ILT should continue to pursue those reliability and validity analyses that have already been conducted with the current versions of the instrument. Specifically, validity evidence should continually be gathered in hopes of strengthening the conviction one has about the inferences made about ILT scores. This will require more standardized administrations of the ILT across a variety of institutions. Other institutions are encouraged to conduct their own studies of the ILT utility and efficacy. We would be pleased to hear about the continuing work conducted at other institutions. Also, item-analyses need to continue to be conducted with the ILT to determine how well the items are functioning.

## 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.
- ACRL (2003), "Guidelines for instruction programs in academic libraries", *College & Research Libraries News*, Vol. 64 No.9, pp.616-9.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Information literacy competency standards for higher education, 2000. Chicago: Association of College & Research Libraries.
- Lewis, D. M., Green, D. R., Mitzel, H. C., Baum, K., & Patz, R. J. (April, 1998). The Bookmark standard setting procedure: Methodology and recent implementations. Paper presented at the annual Meeting of the National Council for Measurement in Education, San Diego, CA.
- Miller, B. J. (2004). Person fit and examinee motivation in a low-stakes testing environment. Unpublished manuscript, James Madison University.
- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24, 3-13.
- Wise, S. L., & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-cased tests. *Applied Measurement in Education*, 16, 163-183.
- Wise, S. L., & Yang, S. (2003). The Adaptex Assessment System (Version 1.5) [Computer Software]. Washington, DC: U.S. Department of Education Fund for the Improvement of Postsecondary Education.

## Section 8. Appendix

### Troubleshooting guide for issues arising in computer-based test administration

It is recommended that all institutions test the URL provided prior to live student testing. Make sure the ENTIRE URL string is copied properly.

If a student has difficulty logging in, please close the window being used and reopen in a new window.

If the program crashes or encounters any problems, please contact technical assistance at +1-303-956-6354.

### Notes for proctors

**Students should not run any programs before or during the test.** As the students arrive, please ask them to take a seat at a computer but DO NOT let them play on the computers. Verify student identity via picture id and do not let test takers use cell phones once they have entered the testing area.