

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

## **Test Manual**

AUTHORS: Steven L. Wise  
Lynn Cameron  
Sheng-Ta Yang  
Susan L. Davis

Revised December 2009  
by Javarro Russell



THE CENTER FOR  
**ASSESSMENT &  
RESEARCH**  
STUDIES

MSC 6806  
Harrisonburg, VA 22807  
540.568.6706 Phone  
540.568.7878 Fax  
assessment@jmu.edu  
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>
<i>2.1 Appropriate and inappropriate uses and interpretations</i>	3
<i>2.2 Target population</i>	3
<i>2.3 Qualifications of users</i>	3
<b>Section 3. Test Development</b>	<b>3</b>
<i>3.1 Academic and theoretical basis</i>	3
<i>3.2 Item type selection</i>	4
<i>3.3 Item pool and scale development process described</i>	4
<b>Section 4. Administrative Procedures</b>	<b>5</b>
<i>4.1 Proctor qualifications and training</i>	5
<i>4.2 Testing procedures</i>	5
<i>4.2.1. Security</i>	5
<i>4.2.2. Process of Administration</i>	5
<b>Section 5. Technical Information</b>	<b>6</b>
<i>5.1 Scoring and interpretation</i>	6
<i>5.2 Evidence of reliability</i>	7
<i>5.3 Evidence of validity</i>	8
<i>5.4 Proficiency Level Standard Setting</i>	9
<b>Section 6. Additional Information</b>	<b>11</b>
<i>6.1 Where to get additional information</i>	11
<i>6.2 Work to be conducted to enhance test use and interpretation</i>	11
<b>Section 7. References</b>	<b>11</b>
<b>Section 8. Appendix</b>	<b>12</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 (Cameron, Wise, & Lottridge, 2007). The ILT is based on the Association of College & Research Libraries (ACRL) Information Literacy Competency Standards (See <http://www.ala.org/ala/mgrps/divs/acrl/standards/informationliteracycompetency.cfm>). 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. The ILT was developed for use at the programmatic level (Cameron et. al., 2007). 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 (Cameron et. al., 2007). 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 (Cameron et. al., 2007). 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 the 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 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 (Cameron et. al., 2007). To determine if the ILT is appropriate for any 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 as appropriate for that particular population.

#### 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 (Cameron et. al., 2007). 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 assesses 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 assesses 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 (Cameron et. al., 2007). 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 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. 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).

It was planned that the ILT would measure the five ACRL standards. However, one of the standards was incompatible with a multiple-choice item format. Standard Four, which refers to the student's being able to use information effectively to accomplish a specific purpose, concerns skills that would be more reasonably assessed through an examination of products or performances that the student produced. Because of this constraint, the ILT items were developed to measure Standards One, Two, Three, and Five. In addition, it was judged that Standards Two and Three should receive greater emphasis on the test; consequently, it was decided that these two standards would each be measured by approximately one third of the test, with the remaining standards comprising 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 freshmen 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.

The final version of the ILT contains 65 multiple-choice items—the 60 items on the final form and 5 items being pilot tested. The pilot tested items will periodically change when adequate information is collected about them and they are replaced by new pilot items. The final test blueprint outlining the specifications for the ILT is shown in Table 1.

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.	12 20% of test	1, 2, 4, 5, 6, 7, 8,9,10,12, 13, 53
<u>Standard 2</u> : accesses needed information effectively and efficiently	19 32% of test	14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 34
<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	3, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 48, 49, 50, 51, 52, 54
<u>Standard 5</u> : understands many of the ethical, legal, and socio-economic issues surrounding information and information technology.	10 17% of test	56, 57, 58, 59, 60, 61, 62, 63, 64, 65
Total Test	60 100% of test	1-60

\*Table adapted from (Cameron et. al., 2007)

## 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 general procedures to be adhered to during the test administration. During training, proctors should be provided with the standardized instructions to be used in the actual testing session. Instructions for administration are provided in the following section.

### 4.2 Testing procedures

The ILT is currently administered via computer software provided by the Center for Assessment and Research Studies. There are certain technical requirements that must be met before the ILT can be administered on a computer. First, because the ILT is administered via the internet, the administering computers must have an internet connection. Second, the resolution on the computer screen must be set to 1024 x 768 and the font size must be set to *medium* so that test questions can be viewed clearly. Third, the *cookies* option on Internet Explorer must be enabled. Finally, the internet program cannot have any popup killers enabled; this will prohibit the test from working properly.

#### 4.2.1. Security

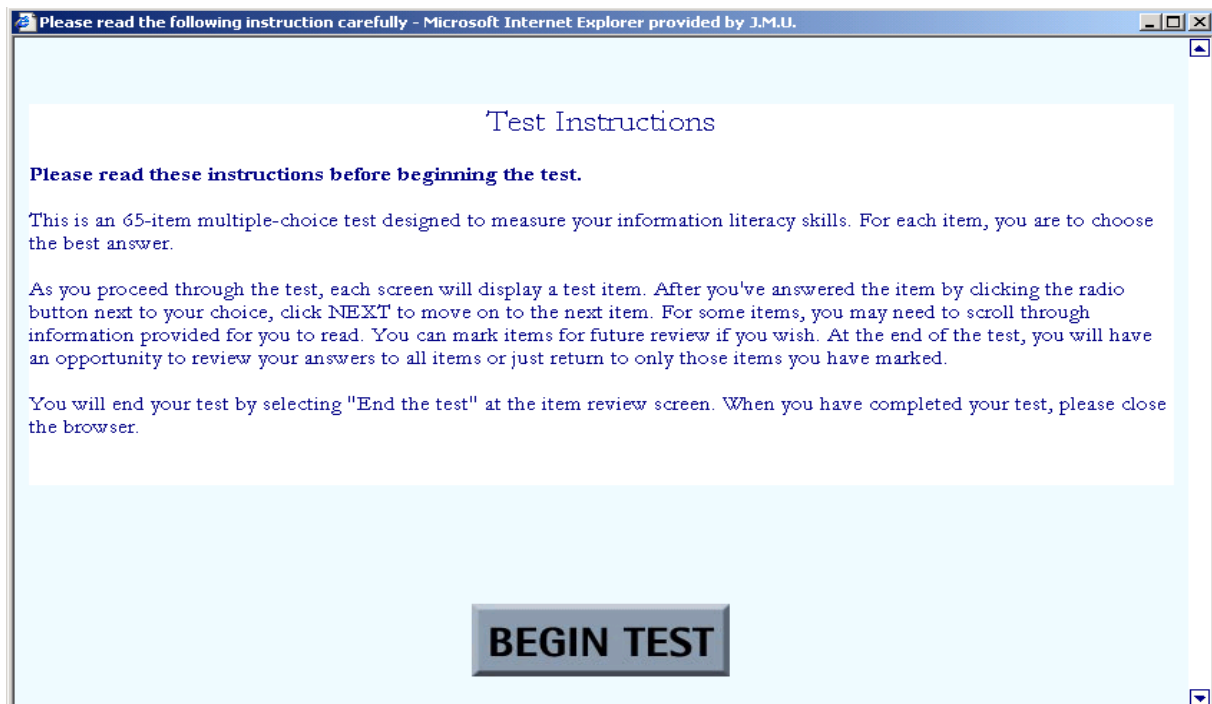
Because the ILT is currently being administered in many different settings, security is of utmost importance. Specifically, caution has been taken to ensure that these items are secure and we ask that test users take the same caution when they have access to the items. The following guidelines will ensure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

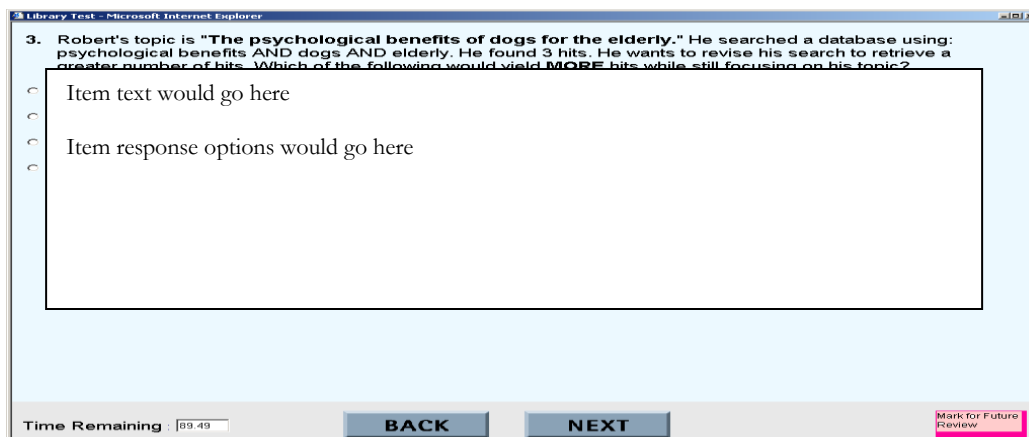
The ILT is a web-based test that is currently supported by the Adaptex web-testing software developed at JMU. Test-takers enter a secure website and provide identifying information and a password to enter the secure software. At the beginning of the test the following testing instructions are presented to students.



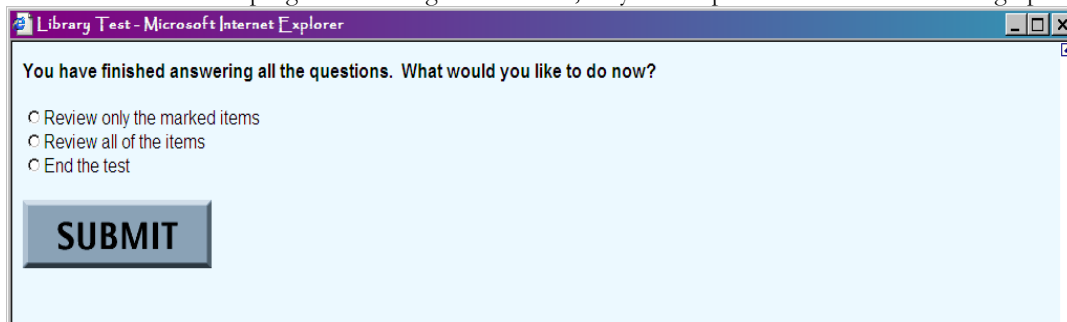
The screenshot shows a web browser window titled "Please read the following instruction carefully - Microsoft Internet Explorer provided by J.M.U.". The main content area has a light blue background and is titled "Test Instructions". Below the title, there is a bold instruction: "Please read these instructions before beginning the test." The text then describes the test: "This is an 65-item multiple-choice test designed to measure your information literacy skills. For each item, you are to choose the best answer." It continues with instructions on how to proceed: "As you proceed through the test, each screen will display a test item. After you've answered the item by clicking the radio button next to your choice, click NEXT to move on to the next item. For some items, you may need to scroll through information provided for you to read. You can mark items for future review if you wish. At the end of the test, you will have an opportunity to review your answers to all items or just return to only those items you have marked." Finally, it says: "You will end your test by selecting 'End the test' at the item review screen. When you have completed your test, please close the browser." At the bottom center of the page, there is a large, dark grey button with the text "BEGIN TEST" in white capital letters.

The ILT in its current version consists of 60 scored items and 5 pilot items. Traditionally, most students are able to complete all 65 items within a 75 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.

Each item is presented on a separate page as shown below (for security purposes, the item is not shown). Each window presents a single item's stem and response options. When students are ready to respond, they should click the radio button next to the desired response option and then click the *Next* button. If students desire to do so, they can select the *Mark for Further Review* option in the lower right hand corner to add the item to a list of those they want to flag for review before completing the test. Before moving to the next item, a response must be given for the current item.



After the examinee has progressed through all 65 items, they will be presented with the following options.



If examinees choose to do so, they can review only those items they marked for further review. A second option is to review all items on the test. The third and final option is to end the test and receive their final scores. Once an examinee submits his or her responses, they are immediately presented with their score on the test (in terms of percentage correct). The testing institution will be provided with a data file containing students scored responses and total scores for each examinee. This data file can either be provided in text format (delimited ASCII file) or in EXCEL format. The scored data files for each institution will be available upon request.

## 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 524 mid-year sophomores in February 2004 during the university's Assessment Day. Table 2 shows, the means, standard deviations, and reliabilities (coefficient alphas) for the total ILT as well as for the items from each of the standards. On average, the sophomore group passed nearly 70% of the items, and exhibited a strong score reliability. The reliability estimates for the standards (subscales), while not as high, are adequate given the parameters for each standard. As noted in the table above, several of these standards are represented by a small number of items which can contribute to 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 on the 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 2. Descriptive statistics and reliability estimates for the 2004 ILT administration at JMU

Scale	# Items	Mean	Standard Deviation	Coefficient Alpha
Total Test	60	41.61	8.45	.88
<u>Standard 1</u> : defines and articulates the nature and extent of information needed.	12	9.70	2.03	.65
<u>Standard 2</u> : accesses needed information effectively and efficiently	19	11.16	2.91	.64
<u>Standard 3</u> : evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system.	19	13.52	3.31	.76
<u>Standard 5</u> : understands many of the ethical, legal, and socio-economic issues surrounding information and information technology.	10	7.18	1.84	.48

N= 524 \*Table adapted from (Cameron et. al., 2007)

Data from year 2008 and 2009 administrations of the ILT were obtained from an aggregated sample of 683 freshmen at four four-year institutions and an aggregated sample of 839 freshmen at five two-year institutions. For the four-year institutions Table 3 illustrates the means, standard deviations, and reliabilities (coefficient alphas) for the total ILT as well as for the items from each of the standards. Table 4 illustrates the same information for the five two-year institutions. On average the freshman students in the four-year and the two-year institution samples passed approximately 60% of the items, and exhibited strong reliability. These reliability estimates are similar to those observed with the JMU sophomore sample in Table 2. As mentioned with the JMU sample, reliability estimates for the subscales are affected by the small number of items associated with each standard. Consequently, subscale scores should only be interpreted at the group or descriptive level. 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 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.

Table 3. Aggregate descriptive statistics and reliability estimates for the ILT at four four-year institutions

Scale	# Items	Mean	Standard Deviation	Coefficient Alpha
Total Test	60	36.12	7.71	.84
<u>Standard 1</u> : defines and articulates the nature and extent of information needed.	12	8.47	2.00	.54
<u>Standard 2</u> : accesses needed information effectively and efficiently	19	8.67	2.53	.54
<u>Standard 3</u> : evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system.	19	12.15	3.04	.69
<u>Standard 5</u> : understands many of the ethical, legal, and socio-economic issues surrounding information and information technology.	10	6.44	1.90	.53

N=683

Table 4. Aggregate descriptive statistics and reliability estimates for the ILT at five two-year institutions

Scale	# Items	Mean	Standard Deviation	Coefficient Alpha
Total Test	60	35.77	7.92	.84
<u>Standard 1</u> : defines and articulates the nature and extent of information needed.	12	8.47	2.08	.58
<u>Standard 2</u> : accesses needed information effectively and efficiently	19	8.28	2.54	.53
<u>Standard 3</u> : evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system.	19	12.14	3.14	.70
<u>Standard 5</u> : understands many of the ethical, legal, and socio-economic issues surrounding information and information technology.	10	6.45	1.89	.50

*N*=839

### 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 ratings of the items were favorable, as all three raters agreed that 42 of the 60 items (70%) matched their intended standard, with at least two raters agreeing that 56 items (93%) matched their standard. Regarding rater agreement, 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 (Cameron et. al., 2007).

Construct validity evidence was obtained through four studies. The first study conducted by (Cameron et. al., 2007) 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 do not try very hard because there are no consequences for 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 freshmen 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 freshmen [ $t(119) = 2.06, p = .041, d = 0.39$ ]. 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 freshmen—collected in fall, 2004—with the scores of 524 mid-year sophomores—collected in spring, 2004. The freshman 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 freshmen, 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.

In a follow-up study, year 2008 and year 2009 ILT data from a sample of 683 freshmen aggregated across four four-year institutions were compared with the ILT scores collected from the sample of 422 JMU freshmen in 2004. The JMU group showed a mean of 37.13 and a standard deviation of 7.70, while the freshman 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 as 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 freshmen aggregated across five two-year institutions were also compared with the year 2004 sample of 422 JMU freshmen. The JMU group showed a mean of 37.13 and a standard deviation of 7.70, while the freshman 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 is 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 really 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 freshmen who had not experienced JMU's information literacy experiences. It was also expected that freshman attending a more selective university, JMU, would perform better than freshman 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. The primary issue they 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 5 illustrates the data obtained from the administrations of ILT discussed in this section. 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.

Table 5 Comparison of ILT mean scores across samples.

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

#### 5.4 Proficiency Level Standard Setting

Without an interpretive context, test scores have little meaning. One way in which the ILT could be used is by administering it to students at two different points in time and assessing 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.

(Cameron et. al., 2007) conducted a standard-setting workshop for the 60-item ILT was conducted during March 2004. An abbreviated 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 *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 6 shows the proficiency definitions given to the judges and the resultant performance standards that the panel recommended. For the Proficient designation, the judges recommended a 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.

Table 6. Performance level definitions and performance 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 LLC through the following Web site: <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. Also, item-analyses need to continue to be conducted with the ILT to determine how well the items are functioning. Further analysis of the pilot items needs to occur as well.

## 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.
- Cameron, L. (2004). Assessing Information Literacy. In Ilene F. Rockman (Eds.), *Integrating Information Literacy into the Higher Education Curriculum* (pp. 207-236). San Francisco, CA: Jossey-Bass.
- Cameron, L., Wise, S.L., Lottridge, S.M. (2007), 'The development and validation of the Information Literacy Test. *College and Research Libraries*, 68, 3, 229-236.
- 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.
- Rodriquez, 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.

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

There are several things that are important to be aware of before administering the ILT in any situation. Many of these specifics will be provided by JMU before the testing date (e.g., web address, passwords, etc.). In addition, the proctor should be familiar with some common questions from test takers that may arise during the testing session. Occasionally, the ILT program may freeze (crash). This is not a major problem; a simple recovery procedure is available to recover a student's data:

1. Close the browser if it is still open.
2. Open the Internet Explorer browser and type in the URL that the student was given. Click **CONTINUE**.
3. Fill in the name, ID in the appropriate boxes and select the institution on the login screen. Select an appropriate test and type the password. Click **SUBMIT**.
4. You will be prompted whether or not you want to resume the test you previously did not finish. Click **CONTINUE**. The program will take you to the question at which the crash occurred.

There are also some common problems that the test proctor should be aware of before administering the ILT. These problems, along with the appropriate solutions, are listed below. In addition, if the proctor experiences a problem for which a solution is not provided, they will be given access to a technical support line before administering the test. Technical support will be available at the provided phone numbers for all (pre-arranged) scheduled testing times.

### Common problems

1. *A few students get blank screens.*

Recover the test

2. *Many students get blank/error screens and cannot recover the test.*

First, make sure that students start the test by typing the URL provided. When test administrators add the URLs as Favorites of the Internet Explorer, the different URLs may be registered because of the nature of our program. Students can check which URL is registered by going to *Favorites, Organize Favorites*, then right-clicking on the icon for the site they use in the Internet Explorer. Find *Properties* and check the URL registered there to make sure that it is identical to the one provided. You can find the URL assigned to each college in the L:\VCCS Testing\serverAssignment.xls. (Click the URLforEachCollege tab).

If students have the correct URL in the Favorites, then have all the students log out and restart the server by following the steps in the “locate a server” and “restart servers” section, then have them recover their tests. If the other servers are not getting maxed out, you can transfer some students to one of them. You will find the max number for each server to be able to handle in the “Servers running” section at the top of this document. L:\VCCS Testing\serverAssignment.xls (Click the assignment tab) tells you how many students are currently accessing each server.

3. *Students get “Page cannot be displayed” errors.*

Click the “Back” link to go back to the last question on which they were working. Tell the test administrator to restart the machines between sessions and to make sure that no one runs any other program **before and during taking the test**.

4. *The program does not accept the password provided.*

Make sure the keyboard is not caps-locked

5. *Cannot recover the test (the program has the student start the new test).*

Make sure the first name, the last name, the section number, and the test name are the same as before.

6. *Cannot see the whole screen.*

Change the screen resolution to 1024x768 and the font size in the browser to medium. You can change the screen resolution by right-clicking the desktop, going to properties, settings, and adjust the screen resolution as 1024 x 768. You can change the text size of the browser by going *View* then *Text Size*.

7. *Cannot find the institution in the dropdown box at login.*

Select any institution temporarily so that we will change the institution name for the student later. Please report this problem to the technical support at JMU so the student's score can be properly recorded under the correct institution.

8. *The test disappears after submit.*

Make sure that the pop up blocker is off. It should be under Tools-> popup blocker in IE.

9. *You see the following dialog box:*

“This page contains both secure and nonsecure items.  
Do you want to display the nonsecure items?  
[Yes] [No] [More info]”

Use the following solution in IE >Tools > Internet Options > Security > Custom Level > Miscellaneous > Display mixed content. By default it is set to prompt, reset it to *Enable*.

10. *Test does not advance past item 1.*

Exit the test and open a new IE.

Under IE >Tool > Internet Options > click “Settings...”

Select the option “Automatically” and click OK.

Now try to sign the back on to the test.

## 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. (When students open music files or play on the computers the files they access take up storage in the cache memory, which creates a problem for the program to run efficiently, so NO ONE is allowed to play at any time on the computers.)

**There are two web addresses provided. Please make sure there are approximately equal numbers of students who receive each URL because the server cannot handle too many students.**

**Restart the computers between test sessions to clear out the computer memory.**

If the program crashes or encounters any problems, you may recover the test (See Troubleshooting section in manual)

**Make sure that the screen resolution is set to 1024x768.**

Students should not double-click the Next/Back button, as it may cause a skipped question.