The Development and Validation of the Information Literacy Test

Lynn Cameron, Steven L. Wise, and Susan M. Lottridge

The Information Literacy Test (ILT) was developed to meet the need for a standardized instrument that measures student proficiency regarding the ACRL Information Literacy Competency Standards for Higher Education. The Web-based, multiple-choice ILT measures both lower- and higher-order skills. Evidence is presented that the ILT scores provide reliable and valid measures of information literacy. In addition, a rigorous standard setting method was used to identify score values corresponding to various absolute levels of proficiency. The ILT can be used to help institutions measure student information literacy outcomes and determine the effectiveness of instruction programs.

Information literacy is a set of competencies that provides a foundation for academic coursework, effective job performance, active citizenship, and lifelong learning. The ALA Presidential Committee defined information literacy as the ability to "recognize when information is needed" and then "locate, evaluate and use effectively the needed information."[1] The sheer abundance of information available in the world today can be overwhelming, and not all of it is reliable. Individuals need to become proficient in the set of skills known as information literacy to be able to conduct an efficient search for information, think critically about the value of a particular piece of information, select sources that are high in quality, and then use the information to accomplish a purpose. This set of skills is important to general education, as well as virtually every major offered in higher education. Information literacy competencies appropriate for higher education have been defined in the form of five standards and twenty-two performance indicators by ACRL.[2]

Instruction programs at college and university libraries provide course-related instruction, tutorials, and other interventions to support student development of information literacy skills. Many programs encourage faculty/librarian collaboration with the goal of helping students develop these skills. In a growing number of institutions, information literacy is formally integrated into the curricula of general education and the majors.[3] With so much interest and em-

Lynn Cameron is Coordinator of Library Instruction at James Madison University Libraries, James Madison University; e-mail: camerosl@jmu.edu. Steven L. Wise is Professor of Graduate Psychology and Coordinator of the Institute for Computer-Based Assessment at the Center for Assessment and Research Studies, James Madison University; e-mail: wisesl@jmu.edu. Susan M. Lottridge is a Doctoral Student at the Center for Assessment and Research Studies, James Madison University; e-mail: daffinsm@jmu.edu.

phasis in both libraries and academic programs, there is a need for reliable and valid data on student learning outcomes. Assessment data can determine the extent to which learning objectives are being met and guide institutions in their efforts to improve curricula and instruction programs.[4] Outcomes assessment can be used to demonstrate the value of the library as a partner in teaching and learning.[5] Assessment of learning outcomes and information literacy has gained prominence in the standards of regional higher education commissions across the country.[6] Despite this growing need and emphasis, and a call by ACRL for institutions to identify assessment methods and develop assessment instruments to measure information literacy outcomes, few standardized instruments exist.[7]

James Madison University (JMU) has a longstanding instruction program that supports student development of information literacy skills at two levels. Students learn basic skills in introductory General Education courses, and they learn skills appropriate to their major field of study later in their academic coursework. The competency-based General Education program specifies that students will be able to "demonstrate competency in information literacy."[8] In the mid-1990s, librarians and General Education faculty wrote specific learning objectives to define this broad goal. To support the broad goal and to help students meet the specific learning objectives, the library developed *Go for the Gold*, a set of Web-based instruction modules that all first-year students are required to complete. In addition, General Education faculty make course-related assignments that require students to find, evaluate, and use information related to course goals. Students demonstrate mastery of basic information literacy skills during their first year by passing the Web-based Information-Seeking Skills Test (ISST).[9] The ISST was developed as a collaborative effort by JMU Libraries and the Center for Assessment and Research Studies (CARS). Librarians used their subject knowledge to write the items, and assessment specialists provided psychometric expertise and advice on test development. Initially, the ISST was used as an assessment test on a large random sample of sophomores. The results were used to measure the effectiveness of the instructional program for first-year students and show where efforts needed to be strengthened. The results were also used in the Southern Association of Colleges and Schools accreditation review and to meet systemwide goals set by the State Council of Higher Education in Virginia.[10] Although the assessment data that had been collected were useful for program evaluation purposes, the university decided that basic information literacy was such an important foundation for academic coursework and lifelong learning that students would be required to pass the ISST to demonstrate competency during their first year at JMU. In 1999, the ISST became a high-stakes test that all first-year and transfer students must pass to continue their studies at JMU. Assessment specialists led faculty and librarians through a rigorous process for setting a standard for passing.[11] The ISST, administered in a secure testing lab, has multiple forms to ensure validity. The scores for the forms are standardized. Success in passing the ISST is noted on the student's transcript.

JMU Libraries supports student learning of discipline-specific information literacy skills by providing each academic department with a liaison librarian who performs course-related instruction for courses in the major and collaborates with faculty to ensure that students develop information literacy skills. Liaison librarians and faculty in a dozen or so departments have written information literacy objectives for the major and developed information literacy assessment tests specific to the discipline. Though not high-stakes tests, these instruments are administered to seniors on a universitywide Assessment Day each spring to determine whether programs are meeting

their goals. The results have been quite useful in academic program reviews, which are conducted every five years, and in accreditation of programs by organizations in the field.

Like JMU, many institutions are being held accountable by governing bodies and professional accreditation boards for demonstrating student outcomes. We have received numerous requests from other institutions to use the ISST to meet their assessment needs. Due to the secure nature of the ISST, we have been unable to provide it to other institutions. We recognized, however, that the need for measures of information literacy extended beyond JMU, which led us to plan a new instrument that could be used by other institutions. In 2002, the JMU Libraries and CARS began to develop the Information Literacy Test (ILT), a new test that could be used by other institutions.

Development of the ILT

Several criteria guided the development of the ILT. To make this new test more generally attractive to other institutions, the ILT was developed to measure the ACRL Information Literacy Competency Standards for Higher Education. It was to be a multiple-choice test that could be completed within one hour. The items were to make frequent use of graphics, documents, and Web page images, and it was anticipated that approximately two-thirds of the items would measure lower-order skills with the remaining 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).

We planned for the ILT to measure all 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, we judged that Standards Two and Three should receive greater emphasis on the test; consequently, we decided that these two standards would each be measured by approximately one-third of the test, with Standards One and Four

TABLE 1
Descriptive Statistics and Reliability Estimates For The ILT

Scale	# Items	Mean	Standard Deviation	Coefficient Alpha
Total Test	60	41.61	8.45	.88
Standard 1: determines the nature and extent of information needed.	12	9.70	2.03	.65
Standard 2: accesses needed information effectively and efficiently.	19	11.16	2.91	.64
Standard 3: 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
Standard 5: understands many of the economic, legal, and social issues surrounding the use of information and accesses and uses information ethically and legally	10	7.18	1.84	.48

comprising the remaining third in roughly equal proportions.

During the first half of 2003, several university reference librarians developed and revised the 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[12] and administered to a sample of 506 incoming freshmen at a medium-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 revised 60-item ILT was subsequently administered to a random sample of 524 midyear sophomores in February 2004 during the university's Assessment Day. Table 1 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 sets of items measuring the individual standards showed the somewhat lower levels of reliability that is characteristic of shorter tests.

Evidence of ILT Validity

Validity concerns the degree to which inferences about students based on their test scores are warranted. The validity evidence for a particular test is typically composed of an accumulation of evidence. For the ILT, two types of evidence have been collected: that based on expert ratings of the items (content validity evidence) 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 evidence).

To assess content validity, three university reference librarians were provided descriptions of the four ACRL standards measured by the ILT and which 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 interrater agreement, all three librarians provided the same rating on 42 of the 60 items (70%), and at least two provided the same rating on 59 items (98%). These results indicate that ILT scores have substantial content validity.

Construct validity evidence was obtained through three studies. The first used the data from the initial administration of the final ILT to university sophomores (spring, 2004 Assessment Day) described earlier. Most of those students had taken the ISST as freshmen. In a study of ISST scores for 333 students, Miller found the correlation between the ILT and the ISST to be significant [$r(331) = .38$, $p < .001$].[13] A problem with data collected on Assessment Day, however, is that some students do not try very hard because there are no consequences for test performance. Wise and Kong showed that item response times can provide a valid measure of the amount of effort a student devotes to a computer-based test.[14] Specifically, Wise and Kong showed that rapid-guessing behavior—in which a response is given more quickly than it would have reasonably taken a student to read a given item, understand its challenge, and identify the correct answer—provides a valid indicator that the student did not give effort to the item. Thirty-six students exhibited rapid-guessing behavior on at least 5% of their items; if these students were deleted from the sample, the ILT-ISST correlation increased to .45.

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 of whom were sophomores. Immediately after taking the ILT, students completed 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 (mean = 40.89, SD = 5.98) scored significantly higher on the ILT than the freshmen (mean = 38.53, SD = 6.17). This difference was statistically significant [$t(119) = 2.06$, $p = .041$, $d = 0.39$]. In addition, ILT scores were significantly correlated with cumulative GPA [$r(119) = .20$, $p = .032$].

The analysis of the survey items revealed positive correlations between ILT scores of the three items concerning student confidence in information literacy. Two of these correlations were statistically significant: "confidence in ability to find books and scholarly articles for projects of interest" [$r(119) = .33$, $p < .001$], and "confidence in ability to evaluate resources for their quality" [$r(119) = .26$, $p = .005$]. The remaining five survey items were more behavioral in nature, addressing how often students visited the library, used the online library catalog, completed *Go for the Gold* instructional modules, and so on. ILT scores were uncorrelated with each of these items, which suggests that although students' information literacy proficiency is related to their self-perceptions of proficiency, such perceptions appear to be unrelated to the number of library-related assignments and activities they had experienced.

The third study compared the ILT scores of 422 incoming freshmen (collected in fall 2004) with the scores of 524 midyear 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.

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.

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 the material. 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 then study the items and 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. An abbreviated version of the Bookmark standard setting method[15] 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

TABLE 2
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%)	**The student who is Proficient is able to:**
		Describe how libraries are organized.
		Define major library services.
		Choose the appropriate type of reference source for a particular information need.
		Identify common types of citations.
		Employ basic database search strategies.
		Locate a variety of sources in a library or online.
		Discriminate between scholarly and popular publications.
		Legally and ethically use information.
Advanced	54 (90%)	**The student who is Advanced is able to attain the criteria for Proficient and:**
		Modify and improve database search strategies to retrieve better results.
		Employ sophisticated database search strategies.
		Interpret information in a variety of sources.
		Evaluate information in terms of purpose, authority and reliability.
		Understand ethical, legal, and socioeconomic issues relating to information access and use.

from data from the fall 2003 pilot testing of incoming JMU first-year students. A diverse panel of ten judges was selected for the workshop: three librarians from our university, three librarians from community colleges in the state, one librarian from another state university, two CARS faculty members, and one doctoral student in assessment.

Two performance standards were set. The first standard differentiated examinees who were *Proficient* from those who were *Below Proficient*. The second one differentiated those who were *Advanced* from those who 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 2 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 clear interpretive benchmarks regarding how many students demonstrated adequate

levels of proficiency in a particular sample.

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.

Summary

Higher education institutions face a growing need to assess the information literacy skills of their students. The purpose of this project was to develop a new Web-based measure of information literacy that can be used by other institutions. The ILT, which can be administered in one hour, measures four of the five ACRL standards. It has been shown to provide reliable scores, we have identified performance standards that may enhance the interpretability of ILT scores, and we have found substantial evidence that its scores from JMU students exhibit validity.

Institutions adopting the ILT, however, should collect their own evidence of score reliability and validity. As with any standardized instrument, it is difficult to judge how the psychometric characteristics of a test at one institution will generalize to other settings and student populations. Thus, while it is likely encouraging to know that evidence for ILT reliability and validity has been found at JMU, it would be prudent for each institution to seek its own evidence.

The ILT is administered via the Internet from secure servers at our institution. Individuals interested in finding out more about the ILT should either contact one of the first two authors or visit the ILT Web site at the following URL: www.jmu.edu/icba/prodserv/instruments_ilt.htm.

Notes

1. American Library Association, Presidential Committee on Information Literacy. *Final Report* (Chicago: ALA, 1989) Available online from www.ala.org/ala/acrl/acrlpubs/whitepapers/presidential.htm. [Accessed 5 April 2007].

2. Association of College and Research Libraries, *Information Literacy Competency Standards for Higher Education.* (Chicago: ACRL, 2000).

3. Ilene F. Rockman, "Strengthening Connections Between Information, General Education, and Assessment Efforts," *Library Trends*, 51, no. 2 (Fall 2002): 185–98.

4. Ibid.; Kenneth R. Smith, "New Roles and Responsibilities for the University Library: Advancing Student Learning through Outcomes Assessment," *Journal of Library Administration* 35, no. 4 (2001): 29–36.

5. Robert E. Dugan and Peter Hernon, "Outcomes Assessment: Not Synonymous with Inputs and Outputs," *Journal of Academic Librarianship* 28, no. 6 (Nov. 2002): 376–80.

6. Bonnie Gratch-Lindauer, "Comparing the Regional Accreditation Standards: Outcomes Assessment and Other Trends," *Journal of Academic Librarianship* 28, no. 1 (Jan. –Mar. 2002): 14–25.

7. Association of College and Research Libraries, *Information Literacy Competency Standards for Higher Education*, <<publishing info?>> 6–7; Elizabeth F. Avery, "Assessing Information Literacy Instruction," in *Assessing Student Learning Outcomes for Information Literacy Instruction in Academic Institutions*, ed. Elizabeth F. Avery (Chicago: ACRL, 2003), 1–5; Lisa O'Connor, Carolyn J. Radcliff, and Julie A. Gedeon, "Applying Systems Design and Item Response Theory to the Problem of Measuring Information Literacy Skills," *College and Research Libraries* 63, no. 6 (2002): 528–43. After we began our test development project, we became aware of Kent State University's Project SAILS, which is an IMLS-funded project to develop a standardized information literacy test. Project SAILS and the ILT are directed at meeting the same general need.

8. James Madison University, "General Education: The Human Community," in *Undergraduate Catalog* (Harrisonburg, Va.: JMU, 2005): 87.

9. Lynn Cameron and Rebecca Feind, "An Online Competency Test for Information Literacy: Development, Implementation, and Results," in *Crossing the Divide: Proceedings of the Tenth National Conference of the Association of College and Research Libraries, March 15–18, 2001, Denver, Colorado*, ed. Hugh A. Thompson (Chicago: ACRL, 2001), 213–18; Lynn Cameron, "Assessing Informa-

tion Literacy," in *Integrating Information Literacy into the Higher Education Curriculum,* ed. Ilene F. Rockman, (San Francisco: Jossey-Bass, 2004), 207–36; Christine E. DeMars, Lynn Cameron, and T. Dary Erwin, "Information Literacy as Foundational: Determining Competence," *JGE: The Journal of General Education* 52, no. 4 (2003): 253–65.

10. State Council for Higher Education in Virginia, *Advancing Virginia through Higher Education: The Systemwide Strategic Plan for Higher Education* (Richmond, Va.: SCHEV, 2002). Available online from www.schev.edu/Reportstats/SystemwideStrategicPlan12-2003.pdf. [Accessed 5 April 2007].

11. Daniel M. Lewis, Donald R. Green, H. C. Mitzel, K. Baum, and Richard J. Patz, "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, Calif., April 1998).

12. Steven L. Wise and Sheng-ta Yang, *The Adaptex Assessment System (Version 1.5)* (Washington, D.C.: U.S. Department of Education Fund for the Improvement of Postsecondary Education, 2003), computer software.

13. B.J. Miller, "Person Fit and Examinee Motivation in a Low-Stakes Testing Environment" (unpublished manuscript, Center for Assessment and Research Studies, James Madison University, Harrisonburg, Va., 2004).

14. Steven L. Wise and Xiaojing Kong, "Response Time Effort: A New Measure of Examinee Motivation in Computer-Cased Tests," *Applied Measurement in Education* 18, no. 2 (2005): 163–83.

15. Lewis et al., "Bookmark Standard Setting Procedure."