A Multi-Method Information Literacy Assessment Program: Foundation and Early Results

William H. Walters, Sarah E. Sheehan, Amy E. Handfield, Bernadette M. López-Fitzsimmons, Susanne Markgren, and Laurin Paradise

abstract: The information literacy (IL) assessment program at Manhattan College in Riverdale, New York, instituted in 2014–2015, evaluates students' information literacy capabilities as demonstrated in their written coursework, their test performance, and their comments on library instruction sessions. Both instruction and assessment are closely linked to five learning objectives, and five years' assessment results have led to significant changes in the IL instruction program. This paper presents key concepts in IL assessment; highlights the importance of evidence-based measures (that is, direct assessment of cognitive outcomes); describes Manhattan College's three assessment methods, with guidelines for the reporting of results; discusses sampling difficulties and related statistical issues; describes the changes in IL instruction undertaken in response to the results ("closing the loop"); and reviews additional assessment methods that can help demonstrate the impact of IL instruction on broader educational outcomes.

Introduction

Formal assessments of educational outcomes have grown in importance over the past two decades. For instance, the number of assessment-related articles and documents indexed by the Education Resources Information Center (ERIC) rose from 514 in 2003 to 1,078 in 2017.[1] Although some faculty and librarians associate assessment primarily with accreditation requirements, information literacy (IL) assessment has the potential to improve teaching and learning in several ways. For example, IL assessment can be used to

1. Determine the extent to which the IL instruction program has been effective;
2. Identify the areas in which students' performance is satisfactory or unsatisfactory;
3. Align instruction more closely with the IL learning objectives;
4. Evaluate whether changes in the instructional program have been effective;
5. Evaluate differences in the performance of first-year students and graduating seniors to gauge the overall impact of the undergraduate experience;
6. Improve the assessment process itself by identifying practices or instruments (for example, rubrics or test questions) that are useful, or not useful, for assessment;
7. Evaluate the effectiveness of instructional methods and materials, such as in-class instruction, library research guides, and online videos;
8. Assess the performance of individual students, targeting those who have the most trouble with some learning objectives; and
9. Examine how IL competencies are associated with broader educational outcomes, such as retention and academic performance.

This paper describes the information literacy assessment program at Manhattan College in Riverdale, New York. The program emphasizes the evaluation of students' written coursework, their IL test performance, and their comments on in-class library instruction. Although several descriptions of multi-method IL assessment programs have appeared in the literature,[2] this paper is noteworthy in several respects. It covers fundamental issues, such as the distinction between cognitive and affective outcomes; it emphasizes evidence-based rather than perception-based measures; it presents a rubric grounded in IL learning objectives broader than those adopted by many universities; it discusses sampling and other statistical issues; it suggests guidelines for the presentation of information in assessment reports; it demonstrates how assessment results can be used to improve instruction; and it highlights the problems likely to arise when implementing an IL assessment program.

Context and Previous Work

Key Concepts

Assessment is most useful when it is closely linked to teaching and learning. In the college or university setting, at least four stages of the learning and assessment cycle can be readily identified:

> **Assessment is most useful when it is closely linked to teaching and learning.**

1. Identify a learning goal or goals consistent with the broader aims of the institution, along with measurable learning objectives that support those goals;
2. Develop and implement an instructional program that helps students achieve the learning objectives;
3. Assess the extent to which the learning objectives have been met; and
4. Modify the curriculum, the instructional methods, or the order or timing of learning activities in response to the assessment results.[3]

The fourth stage, sometimes called "closing the loop," is intended to improve instruction, but it can also involve modification of the learning objectives or assessment activities. Once the fourth stage has been completed, the cycle repeats, so the program can again be evaluated and further improved.

Most guides to educational assessment describe three important distinctions.[4] The first is the distinction between *inputs* and *outcomes*. Library assessment efforts have historically relied on the reporting of inputs, such as the number of books acquired and the number of instruction sessions offered. Input data are still routinely collected by such agencies as the Association of College and Research Libraries (ACRL) and the National Center for Education Statistics.[5] Inputs do not necessarily coincide with students' knowledge or performance, however, and recent decades have seen a major shift from the reporting of inputs to the use of outcome data, such as students' test scores.[6]

A second distinction can be made between *cognitive* and *affective* measures or outcomes, which correspond to cognitive and affective learning domains.[7] Cognitive measures are appropriate whenever the goal is to assess students' learning or performance, such as their ability to remember, understand, apply, analyze, evaluate, or create. In contrast, affective measures are associated with students' perceptions, values, and motivations. A cognitive measure might evaluate the appropriateness of the information resources that students use in their work, for instance, while an affective measure might evaluate their engagement, confidence, or comfort with the process of evaluating information. Many affective measures deal with self-efficacy, an individual's belief in his or her ability to succeed at an endeavor.

A third distinction can be made between *direct* measures, which are closely linked to cognitive or affective constructs, and *indirect* measures, which evaluate those same constructs through more roundabout methods. Although some authors describe cognitive measures as direct measures and affective measures as indirect,[8] the distinction is more nuanced than that. For instance, if we want to measure students' cognitive abilities regarding Task X, we can administer a test (direct method) or ask a survey question, "How good are you at Task X?" (indirect method). As a cognitive measure, the survey question is less helpful because it will more likely generate biased or uninformed responses. However, that same survey question can be regarded as a direct measure—and a good one—if the goal is not to assess students' abilities but to evaluate their confidence in their abilities. The question gauges affect directly but ability only indirectly.

Current Practices

The ACRL Assessment in Action site provides brief summaries of the IL assessment initiatives at 188 North American colleges and universities.[9] Summary data for the 38 institutions most like Manhattan College—master's institutions with full-time equivalent enrollments of 2,000 to 9,999—can be seen in Table 1.[10] Notably, none of the 38 assessment programs make much use of input data; all focus on outcomes rather than inputs.

As Table 1 shows, 87 percent of the assessment programs evaluate one or more cognitive outcomes. Just 21 percent measure affective outcomes. The single most common assessment activity is the direct assessment of students' written work—the use of a rubric to evaluate research papers.[11] Testing is less commonly undertaken, as are the

Table 1.

Characteristics of assessment programs: ACRL Assessment in Action versus Manhattan College

Characteristic	ACRL Assessment in Action* percentage with yes responses	Manhattan College Assessment of written work	Manhattan College JOLT/BOLT IL test†	Manhattan College Student evaluations of IL instruction
IL is an institutional core competency	71	yes	yes	yes
Cognitive and/or affective measures (outcomes)				
Cognitive: e.g., students' learning	87	yes	yes	yes
Affective: e.g., students' engagement	21	no	no	no
Assessment methods and data				
Direct: assessment of students' written work	79	yes	no	no
Direct: test performance	47	no	yes	no
Indirect: student surveys or self-evaluations	49	no	no	yes
Indirect: student interviews or focus groups	16	no	no	no
Indirect: instructor observations or reflections	13	no	no	no
Indirect: use of library collections or services	3	no	no	no
Multiple assessment methods	71	—	—	—
Population: educational level				
Undergraduates	84	yes	yes	yes
First-year students	13	yes	yes	yes
Seniors	0	yes	yes	no
Graduate students	5	no	no	no

Population: subject areas				
Broad range of subject areas	66	yes	yes	no
English composition	34	yes	no	yes
Other subjects (3 or fewer)	13	no	no	no
Linked to data on student success	24	no	no	no

*ACRL Assessment in Action refers to the 38 participating master's colleges and universities with full-time equivalent enrollment of 2,000 to 9,999. When taken by seniors, JOLT is known as BOLT.

†JOLT (Jasper Online Information Literacy Test) is Manhattan College's first-year test of IL proficiency.

various indirect methods of assessment. In fact, only one indirect method—surveys or self-evaluations—was adopted by more than 16 percent of the institutions in the sample. (Like the Assessment in Action staff, the authors assume that survey questions and interviews are indirect methods of assessing cognitive outcomes rather than direct means of evaluating affective outcomes.) More than 70 percent of the institutions represented in Table 1 used multiple assessment methods. Assessment data are generally more reliable, and more convincing, if multiple techniques yield similar results.[12]

> **Assessment data are generally more reliable, and more convincing, if multiple techniques yield similar results.**

Surprisingly, none of the 38 colleges and universities have singled out graduating seniors for assessment. Although many accreditors—and, presumably, many employers—are interested in students' capabilities when they enter the professional job market,[13] senior-year assessment seems not a priority for Assessment in Action participants.

As shown in Table 1, most IL assessment programs include students in a wide range of subject areas; they evaluate the capabilities of the entire student body. When assessment is concentrated within just one subject area, that area is nearly always English composition. In describing their programs, several Assessment in Action participants noted that IL assessment and writing assessment are coordinated in one way or another.

Nearly a quarter of the assessment programs represented in Table 1 have explicitly considered the links between library-related variables (such as frequency of library use) and broader indicators of student success, such as grades, retention, graduation rates, and job placement. This approach has great potential for demonstrating the impact of the library on students' educational outcomes.[14]

The Importance of Evidence-Based Assessment

The IL assessment literature suggests that evidence-based measures, which require students to demonstrate their knowledge or skills, are generally superior to perception-based measures, which ask them to report on their opinions or their perceived capabilities. The distinction between evidence-based and perception-based measures is nearly identical to the distinction between cognitive and affective measures.[15]

Perception-based measures feature heavily in several standard assessment instruments. For instance, the ACRL LibQUAL+ survey asks students to indicate their satisfaction with various aspects of the library's services, collections, and facilities.[16] Likewise, the IL-HUMASS (Information Literacy Humanities and Social Sciences) survey asks students to rate their own abilities on several IL tasks in each of four categories: information search, information evaluation, information processing, and information communication and dissemination.[17] The National Survey of Student Engagement (NSSE) also includes five survey questions that can be mapped to standard IL competencies. The NSSE questions do not assess students' performance at all, however; they simply count how often students have undertaken various activities, such as working on collaborative projects and discussing ideas from their course readings.[18] Perception-based measures also feature prominently in many locally developed IL assessment instruments.[19]

Although perception-based measures are appropriately used to gauge attitudes or opinions, they are sometimes interpreted as indicators of ability or effectiveness. For instance, the LibQUAL+ survey and a recent guide to library assessment are each based on the assumption that customers' perceptions are the only valid measures of service quality.[20] There are several problems with this approach, however.[21] Among other things, we can question:

1. Whether students' self-assessments are in line with their demonstrated capabilities;
2. Whether students have the knowledge and expertise needed to evaluate the quality of instruction;
3. Whether their assessments are biased by extraneous factors, such as the instructor's gender;[22]
4. Whether students have the range of experiences needed to distinguish between effective and less effective library programs and services;
5. Whether students' preferences and wants (which often focus on performance goals) are fully aligned with their needs (the learning goals established by the instructor and the institution);[23]
6. Whether perceptions can ever serve as valid indicators of objective conditions; and
7. Whether any client can judge professional services in more than a superficial way, since the professions are generally defined as the occupations for which clients are not qualified to judge the practitioner's competence.[24]

The first point is especially important, since empirical research has shown conclusively that students (especially students with poor information literacy skills) tend to overestimate their own IL abilities.[25] For most components of information literacy, students' self-assessment scores are no better predictors of ability than the values that might be produced by a random

> … empirical research has shown conclusively that students (especially students with poor information literacy skills) tend to overestimate their own IL abilities.

number generator.[26] In fact, self-assessment scores are generally better indicators of self-efficacy or self-confidence than of abilities.[27] It is important to acknowledge students' perceptions, especially since individuals with favorable attitudes will more likely use the library and seek assistance.[28] Nonetheless, most surveys, interviews, and focus groups evaluate patrons' opinions rather than their abilities.

As Table 1 shows, tests are a useful evidence-based method of assessing cognitive IL outcomes. Two standardized tests, the Standardized Assessment of Information Literacy Skills (SAILS) and the Madison Assessment Information Literacy Test, are currently in widespread use. Both present students with approximately 60 multiple-choice questions that address four of the five IL standards set forth by ACRL in its Information Literacy Competency Standards for Higher Education.[29] Other cognitive IL tests include the Research Readiness Self-Assessment and the Beile Test of Information Literacy for Education.[30] Each of these tests was developed and validated with great care, but all four

are keyed to the ACRL Standards and may not incorporate all the learning objectives established by any university. Locally developed instruments, such as brief exercises and quizzes, can provide a more nuanced and sometimes more appropriate alternative to standardized tests.[31]

> **Locally developed instruments, such as brief exercises and quizzes, can provide a more nuanced and sometimes more appropriate alternative to standardized tests.**

Although tests remain important, they have been overshadowed in recent years by the evaluation of students' research papers, bibliographies, assignments, or portfolios. This shift reflects a trend toward authentic assessment, in which students demonstrate their abilities in situations much like those they will encounter in the classroom and the workplace.[32] Authentic assessment usually involves a rubric or scoring sheet grounded in the institution's IL learning objectives. Librarians or faculty evaluate students' written work, assigning scores that can be used to determine the percentage of students who meet the minimum standard in each area. The ACRL rubric is widely used, although other IL rubrics may be more appropriate for some settings.[33] The Manhattan College rubric is presented in the Appendix.

Information Literacy at Manhattan College

The Manhattan College Context

Manhattan College is a private master's university with 3,500 undergraduates and 500 graduate students. The college is a member of the Oberlin Group, a consortium of leading liberal arts colleges, and is distinctive in its emphasis on engineering and business as well as the arts and sciences. O'Malley Library, which is open 24 hours a day, 7 days a week during the academic year, provides IL instruction in a setting that is strongly residential and student-centered.

There are two components to library assessment at Manhattan College. First, the library participates in the college-wide assessment process for administrative units. This process, based on the library's annual goals, does not normally include IL but does include periodic evaluations of access services, technical services, collection development, and other operational areas.[34] The second component of assessment, described in this paper, is the library's IL assessment program, which focuses exclusively on outcomes—students' performance—rather than inputs such as the number of library instruction sessions offered. Specifically, the program (1) examines the written work that students complete for their courses; (2) evaluates their performance on two online information literacy tests, the JOLT (Jasper Online Information Literacy Test) for first-year students and the BOLT for seniors; and (3) reviews their evaluations of the IL instruction sessions. The IL assessment program is especially useful for accreditation, since it demonstrates that the library undertakes systematic assessments and responds to the results when allocating resources and planning instruction.

IL Learning Objectives

At Manhattan College, information literacy is one of the seven core competencies expected of all undergraduates, regardless of major. The *core competency* designation signals to internal and external audiences that IL is taken seriously by the college administration and by the regional accreditor. The library has responsibility for students' success in IL, and it is the only core competency for which primary responsibility lies outside the academic departments.

The college's IL learning goal, to ensure that graduates can "evaluate and select appropriate information resources, integrate them into original work, and cite them correctly," is more fully represented by five learning objectives, designated LO1 to LO5 (see Table 2). The learning objectives are based on the ACRL Information Literacy Competency Standards, the 2012 ACRL Guidelines, and the 2013 Information Literacy VALUE (Valid Assessment of Learning in Undergraduate Education) Rubric, although they are somewhat more inclusive than the standards set forth in those documents.[35] (See LO4, integrating information into academic work.) The college has not adopted the 2016 ACRL Framework for Information Literacy for Higher Education, however, mainly because of the need to maintain consistency over time in its assessment methods.[36] The Framework would require the revision of both the learning objectives and the assessment rubric, making it difficult to compare students' performance over time. Two institution-specific factors are also relevant. First, the library has worked for five years to make the faculty aware of the current learning objectives, which are widely supported. A change would require reestablishing that support. Second, while the Framework is admirable in its breadth, the Manhattan College librarians have not reached a consensus on its utility as a guiding document for day-to-day IL activities.

Each learning objective is linked to the college's strategic plan and informed by interactions with the Provost's Office and the Institutional Effectiveness Committee. Although some authors have argued that colleges and universities ought to agree on uniform IL standards,[37] Manhattan College maintains that each institution's learning objectives should vary with differences in mission, goals, curricula, resources, and student and faculty populations.

IL Instruction

While the IL instruction program has several components, it is grounded in the classroom instruction provided within the required first-year English composition course. Forty sections of composition are offered each year. Each includes at least one 50- or 75-minute IL instruction session taught by a librarian with the assistance of the course instructor, and about 40 percent of the sections incorporate more than one IL session. All the librarians teach within the composition course, and their methods include lectures, discussions, group activities, informal quizzes and competitions, guided completion of worksheets, and assisted practice in searching. The presentations, notes, and handouts for each session are posted to the instructor's site on Moodle, the college's open-source learning management system. Nearly every session provides students with skills that will be immediately useful for a graded assignment within the composition course. Interaction between the librarian and the course instructor is essential, and librarians

Table 2.
Manhattan College information literacy learning objectives (LOs)

LO1. Identifying information needs

Identify the information needs that correspond to an academic task (e.g., assignment) and the kinds of resources or documents that are likely to meet those needs. This objective refers to the kinds of resources the student chooses—not to the items chosen. Among other things, students must demonstrate an understanding of the differences between scholarly and popular sources and the distinction between opinion and evidence.

LO2. Searching and retrieving documents

Demonstrate an awareness of best practices in the identification and retrieval of relevant books, articles, online resources, video files, and other documents. Among other things, students must be able to identify relevant databases and to construct searches that are likely to be effective.

LO3. Evaluating and selecting documents

Evaluate and choose information resources with regard to their general characteristics (topic relevance, currency, authority, bias, etc.) and their appropriateness for a particular task. Unlike LO1, this objective refers to the specific items (articles, books, etc.) chosen by the student.

LO3 includes 3.1 (relevance of sources to the student's needs—items chosen) and 3.2 (quality of sources—items chosen).

LO4. Integrating information into academic work

Integrate information resources into written or presented work in ways that support the goals of the academic task or assignment (e.g., as context, presentation of competing views, or supporting evidence).

LO4 includes 4.1 (degree to which student has incorporated information resources when it is necessary or appropriate to do so) and 4.2 (task-specific relevance and use of sources).

LO5. Citing sources and using citations

Cite information sources completely, appropriately, and in keeping with the norms of scholarly writing or presentation. Demonstrate understanding of basic intellectual property issues. Among other things, students must be able to avoid plagiarism and misrepresentation.

LO5 includes 5.1 (attribution and identification of cited works) and 5.2 (completeness and format of citations).

offer library instruction only when the instructor is present. The director of the writing program, a professor of English, has provided advice and helped secure the cooperation of the 12 faculty who teach composition each year.

Apart from in-class instruction in composition courses, Manhattan College librarians teach IL through instruction in other courses; independent workshops on such topics as searching, plagiarism, and citing; reference work; the development of online library research guides for subjects; and an online tutorial that helps students prepare for the JOLT information literacy test.[38]

Assessment and Reporting

As noted earlier, the IL assessment program has three main components, which are sum-marized in Table 1. The program (1) examines the written work that students complete for their courses, (2) evaluates their performance on the JOLT and BOLT online information literacy tests, and (3) reviews their evaluations of the IL instruction sessions. Not all as-sessment activities are undertaken every year, however (see Table 3). Manhattan College librarians employ all three of the methods most often used by ACRL Assessment in Action participants, as listed in Table 1. Their program is distinctive, however, in its emphasis on graduating seniors and its incorporation of more than two assessment methods.

Table 4 outlines the content of the first annual information literacy assessment re-port, a 28-page document completed in June 2015. Subsequent reports have been about half that length, focusing almost exclusively on sections three and four ("Assessment activities" and "Closing the loop"). In June 2019, the library completed the fifth annual report, which includes cumulative results for the first five years of assessment. The library will place IL assessment activities on hold for the next two years while it continues to modify instruction in accordance with the first five years' results. Annual assessments will resume in 2021–2022, when the library will evaluate whether its changes in IL in-struction have improved students' performance.
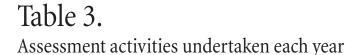
First Assessment Method: Evaluation of Students' Written Work

Assessment Activities

The evaluation of students' written work is central to the Manhattan College assessment efforts. The papers appraised are most often term papers, but they can be any assign-ments that make significant use of information resources and were submitted for a grade in a regular course. In 2014–2015, librarians assessed the papers of first-year composition students. In subsequent years, papers were solicited from the instruc-tors of senior-year courses in all subject areas. The switch from first-year students to seniors was undertaken for three reasons. First, librarians wanted to evalu-

> … senior-year assessment gauges students' capabilities when they enter the profes-sional job market; data compiled shortly before graduation may be especially meaningful to students, parents, potential employers, and accrediting agencies.

Table 3.

Assessment activities undertaken each year

Activity	2014–2015	2015–2016	2016–2017	2017–2018	2018–2019
Written work: first-year students	50	—	—	—	—
Written work: seniors	—	40	23	34	—
JOLT: first-year students*	971	1,366	1,138	1,060	1,043
BOLT: seniors	—	—	—	—	105
Student evaluations: first-year students	561	622	209	365	—

Each number is the sample size—the number of papers evaluated, the number of test takers, or the number of student evaluations received.

*JOLT (Jasper Online Information Literacy Test) is Manhattan College's first-year test of IL proficiency. When taken by seniors, JOLT is known as BOLT.

ate their entire IL instruction program—not just the first year. Second, they wanted to assess long-term skill development and understanding rather than short-term retention. Third, senior-year assessment gauges students' capabilities when they enter the professional job market; data compiled shortly before graduation may be especially meaningful to students, parents, potential employers, and accrediting agencies.[39]

Each year, an assessment team of three librarians evaluated each sampled paper to assess students' performance regarding the college's IL learning objectives. Learning objectives LO1 and LO2 could not be readily assessed using students' written work, so they were excluded. LO3, LO4, and LO5 were each split into two components to provide a clearer sense of students' abilities in the aspects of information literacy. (See the Appendix.) Although the IL criteria are not necessarily related to the standards that instructors use when grading students' papers, the learning objectives are consistent with the academic norms of the college. Librarians also work with faculty to include IL as a graded component of students' work whenever possible.

From 2015–2016 onward, each year's assessment team included at least two librarians who had previously served on the team. Each team member assigned six scores to each paper, one for each of the assessment criteria shown in the Appendix. The four possible scores were 1 (below minimum standard—major problems), 2 (below minimum standard—some problems), 3 (meets minimum standard—essentially proficient), and 4 (exceeds minimum standard—fully proficient). Because these performance levels are clearly defined and mutually exclusive, fractional scores are not permitted. Evaluating students' work over a four-year period, the library developed several principles that supplement the rubric itself:

Table 4.

Outline of the 2014–2015 Manhattan College information literacy assessment report

1. Introduction and context
     The assessment process
     Learning goal and learning objectives
2. Learning activities
     Course-embedded library instruction
     Online tutorial
     Other learning activities
3. Assessment activities, 2014–2015
     Review of students' written work
          Assessment procedure
          Assessment results
     The JOLT*
          Assessment procedure
          Assessment results
     Students' evaluations of library instruction sessions
          Assessment procedure
          Assessment results
4. Closing the loop: Instructional responses to the 2014–2015 results
5. Planned assessment activities, 2015–2016 and subsequent years
     Refinement of assessment instruments and procedures
          Review of students' written work
          The JOLT
          Students' evaluations of library instruction sessions
     Senior-year assessment

The 2014–2015 report also includes three appendices: "Learning objectives represented in the JOLT test," "JOLT test questions and data," and "Rubric for assessing students' written work."
*JOLT (Jasper Online Information Literacy Test) is Manhattan College's first-year test of IL proficiency.

1. Librarians' basic expectations for information resources (for example, currency, reliability, authority, and purpose) are included in criterion 3.2 (quality of sources). In contrast, 3.1 (relevance of sources to the student's needs) evaluates whether the resources are functionally relevant—whether they advance or support the student's argument, for instance.

2. In many cases, nonscholarly works are appropriate for the student's purposes. Engineering students routinely cite manufacturer's catalogs or product specifications, and business students make use of corporate websites and advertisements to discuss the marketing strategies employed by companies. The librarians try to avoid assuming that scholarly resources are especially appropriate.

3. Criterion 4.1 (degree to which the student has incorporated information resources when it is necessary or appropriate to do so) does not refer to students' citing of the information resources they used. Instead, it evaluates whether they made use of information resources when they ought to have done so—when external evidence or theory was obviously relevant to the student's presentation or argument. The librarians relied on their individual judgment regarding the range of information that was potentially available for students to use. They assumed, for instance, that basic statistics on gun violence were available for every European nation, but they did not expect students to find or cite the same statistics for less developed countries.

4. Because the expectations for adherence to any citation format vary considerably from one instructor to another, one of the two citation criteria (5.1, attribution and identification of cited works) is independent of format. Criterion 5.1 does not require any specific citation elements—not even title or author—if the source can be identified.

5. The librarians considered it legitimate to assign low scores for two separate criteria if one of them built on the other. For instance, 5.2 (completeness and format of citations) assumes that the student has cited his or her sources, as specified in 5.1. A student who failed to cite altogether would therefore receive low scores for both 5.1 and 5.2.

Each librarian (rater) on the assessment team spent about 15 minutes scoring each paper. Once the initial scores were obtained, the three raters met as a group to discuss the papers for which a single rater's score was two or more points different from another rater's score as well as the papers that posed special problems regarding the rating criteria. The goal was not to reach a consensus or to change anyone's scores, but to make sure the raters had fully understood the rating criteria. Nonetheless, each year's two-hour meeting did result in score changes for about half the papers discussed. For each paper, the raters calculated the final score for each of the six rubric criteria by taking the median of the three raters' scores.

Reporting of Results

For each year's IL assessment report, the library first presented Cronbach's alpha, a measure of internal consistency (inter-rater reliability) for each of the six criteria shown in the Appendix.[40] For each criterion, the report also showed the percentage of scores for which one rater assigned a final score two or more points different from that of another rater.

In presenting the final scores, the library followed the format shown in Table 5. In some years, it highlighted changes over time by presenting a separate table for each year. Although Table 5 is intended mainly to display the format of the table rather than the results, it shows that Manhattan College students' performance is unsatisfactory for 4.1

Table 5.

Example table: Seniors' scores for each of the six rubric criteria, 2015–2016 through 2017–2018

Criterion	3.1	3.2	4.1	4.2	5.1	5.2
Average score (1–4)	3.3	2.8	2.4	2.5	2.6	2.2
Standard deviation*	0.8	0.9	0.8	0.8	1.0	1.0
% with score of 1	5.2	8.2	16.5	15.5	21.6	29.9
% with score of 2	5.2	27.8	30.9	24.7	14.4	22.7
% with score of 3	45.4	43.3	47.4	57.7	43.3	40.2
% with score of 4	44.3	20.6	5.2	2.1	20.6	7.2
% with score of 3 or 4	92.5	65.9	54.2	61.6	65.9	48.9

*The standard deviation represents the dispersion or variability of the scores.

(degree to which the student has incorporated information resources when it is necessary or appropriate to do so) and 5.2 (completeness and format of citations) but satisfactory for 3.1 (relevance of sources to the student's needs). For each criterion, the average score is less important than the percentage with unsatisfactory scores—scores of 1 or 2.

The library's discussion of the results has varied from year to year, but it generally examines students' overall performance, identifies the areas in which their performance is unsatisfactory, and discusses changes over time, with emphasis on the apparent effects of any instructional changes made in response to previous reports. For areas of unsatisfactory performance, the librarians also present their subjective observations, noting, for instance, that students in some disciplines tend to cite many papers in a casual way without making connections to their own work or explaining why the cited works are relevant.

Sampling Difficulties

As Table 3 shows, three of the annual assessments used the written work submitted for senior-year courses. Manhattan College has no senior thesis requirement, however, and not all majors require writing-intensive courses in the senior year. Each of Manhattan College's six schools is independent in terms of curriculum, and there is currently no reliable list of senior-year courses with a writing component.

The library solicited senior-year papers by working with the deans, the department chairs, the instructors of courses known to be writing-intensive, and the faculty who use the library regularly. The number of papers received from each school varied dramatically from year to year, however. Table 6 shows the most extreme case, from 2017–2018,

when the two largest schools, Engineering and Business, submitted no papers at all. Likewise, the library received papers from 15 of the 30 graduating seniors (50 percent) in the School of Continuing & Professional Studies but from just 8 of the 74 seniors (10.8 percent) in the School of Science. Overall, it collected 57 papers but used only 37 of them. This procedure was necessary to create a sample that was representative of the four schools that submitted papers.

The foremost requirement for reliable results is not sample size, but representativeness (see Table 7).[41] Even without making cross-school comparisons, the library still wanted a college-wide sample that was representative in terms of the six schools within Manhattan College. Since 21 percent of the college's bachelor's graduates are in the School of Liberal Arts, for instance, the library constructed a sample in which 21 percent of the papers came from students in that school. To ensure representativeness, the team first divided the number of papers received by the number of graduating seniors for each of the four schools that submitted papers. The school with the lowest number of papers per graduate (Science) determined the sampling rate (10.8 percent) for all four schools that submitted papers. As Table 6 shows, the resulting sample has 16 papers (152 × 10.8 percent) from the School of Liberal Arts, 10 (97 × 10.8 percent) from the School of Education & Health, 8 (74 × 10.8 percent) from the School of Science, and 3 (30 × 10.8 percent) from the School of Continuing & Professional Studies. Because the schools of Engineering and Business were not represented in the 2017–2018 data, the results for that year cannot be extended to those two schools.

> **The foremost requirement for reliable results is not sample size, but representativeness.**

Intensive efforts to solicit senior-year papers may help the library construct larger, more representative samples in the future. Unfortunately, three problems persist despite the library's best efforts. First, some faculty—including several with a good understanding of sampling, research methods, and assessment—have been reluctant to send the library all their students' work. Some want to submit just the best papers. Second, some faculty do not see the need for IL assessment because "the papers have already been graded." Although the library's assessment criteria are distinct from those used to grade students' papers, several instructors have questioned the need for further evaluation. Finally, some faculty believe it inappropriate to share students' papers with anyone, or with anyone outside the department. The library removes all grades and identifying information before examining the papers, mainly to prevent bias in its own scoring procedures, and has invited the faculty to do so before submitting the papers to the librarians. Still, a small number of instructors are wary about sharing their students' work.

> **… the evaluation of students' written work—the direct, authentic assessment of cognitive outcomes—is the single most useful IL assessment exercise …**

Despite these difficulties, the evaluation of students' written work—the direct, authentic assessment of cognitive outcomes—is the single most useful IL assessment exercise. It evaluates the impact of the library's entire IL program, emphasizes the capabilities that students can demonstrate near the end of their undergraduate careers, and

Table 6.
Sampling procedure for senior-year papers, 2017–2018

School	Papers received sample*	Graduates	Papers per graduate (%)	10.8%
Engineering	0	207	0.0	—
Business	0	160	0.0	—
Liberal Arts	20	152	13.2	16
Education & Health	14	97	14.4	10
Science	8	74	10.8	8
Continuing & Professional Studies	15	30	50.0	3
Total	**57**	**720**	—	**37**

*A 10.8% sampling rate was used to obtain a sample representative of the four schools that submitted papers.

Table 7.
Guidelines for assessment-oriented sampling

1. Representativeness is more important than sample size.
   Although large samples can reduce one particular type of error, sampling error, other errors and biases—such as imprecise delineation of the study population, nonrepresentative sampling, selection bias, and measurement error—often have a greater impact on the reliability of data and results. These same errors and biases can invalidate tests of statistical significance, which account for sampling error only. The assessments of students' written work did not make use of significance tests, since an important requirement of those tests—random or representative sampling—was not met.
2. The size of the population has no bearing on the number of cases or responses needed for a reliable analysis.
   All else equal, the sample size needed to estimate the average IL test score at a university of 30,000 students is no greater than the sample size needed to estimate the average IL test score at a college of 2,000 students. Unlike the size of the sample, the size of the population has no bearing whatsoever on statistical significance or on the reliability of the assessment results.

Table 7, continued.

3. The increase in statistical power that comes with greater sample size diminishes as sample size increases.

    Statistical power is related to sample size, but the relationship is not linear. For instance, an increase in sample size from 40 to 50 will have a far greater effect than an increase in sample size from 400 to 410.

4. The necessary sample size depends on the comparisons the investigator wants to make—the groups to compare.

    In the view of the Manhattan College team, a representative sample of 50 students per year is probably sufficient for the assessment of students' written work. That sample size allows them to compare students' average scores, college-wide, from one year to the next. However, comparing 2017–2018 scores with 2018–2019 scores for each of the six schools of the college would require that same number of cases within each of the groups being compared (6 schools × 50 students each = 300 students per year).

5. If incentives are offered to students who participate, they should be planned in a way that avoids bias, since a large, biased sample is generally worse than a small, unbiased sample.

    All participants should be motivated to do well rather than to complete the exercise with the least amount of effort. If participation is associated with the opportunity to win a prize, the odds of winning should be higher for those who do especially well. At the same time, even students who expect to do poorly should have some chance of winning.

focuses on students' performance—what they can do rather than how they respond to test questions or how they rate their own abilities. The evaluation process also provides a good sense of the subtleties related to students' use of information resources. Because nearly all the librarians have participated in the process, they share an understanding of the ways in which students tend to meet, or not meet, each of the IL learning objectives.

Second Assessment Method: The JOLT and BOLT Tests

Assessment Activities

At Manhattan College, all first-year students must successfully complete the Jasper On-line Information Literacy Test (JOLT), which is designed to cover the IL fundamentals that most students learn in high school. Students may retake the test as many times as necessary, but they must achieve a score of 70 percent or higher before they can register for their second-year courses. Those who fail the test by more than a few points are advised to meet one-on-one with a librarian or to complete an online tutorial before taking it again.

The current JOLT test, hosted on the Moodle learning management system, has 30 questions, each keyed to one or more of the five IL learning objectives (see Table 8). The order of the questions and of the responses are randomized each time the test is taken. Twenty-two of the 30 questions are always presented, so students who take the test more than once will encounter those same questions every time. The other eight questions

Table 8.

Typical Jasper Online Information Literacy Test (JOLT) questions that address the five learning objectives (LOs)

LO1. Identifying information needs

You need information about recent changes in copyright law. Which would be the most appropriate places to start your research? (Choose two.)*

    Search a periodical database
    Call a criminal lawyer
    Search the online catalog of books and DVDs
    Search Google
    Phone a friend

LO2. Searching and retrieving documents

You are searching a library database for articles about college students' jobs. Which of these is the better search strategy?

    College AND students AND work
    College OR students OR work

LO3. Evaluating and selecting documents

Which of the following are NOT recommended criteria for evaluating a website? (Choose two.)

    Author
    Date
    Graphic design
    Publisher
    Font

LO4. Integrating information into academic work

Which of the following are good ways to avoid plagiarism? (Choose all that apply.)

    Keep track of citations (bibliographic information) for all the information sources you use
    Take accurate notes that show where you found specific ideas
    Paraphrase authors' ideas without citing the source in which they were found
    Use quotation marks around exact quotations and cite the source (including the page number)
    Change the font when you copy and paste into your paper to make it clear which text was taken directly from the source

Table 8, continued.

LO5. Citing sources and using citations

If you had this citation for an article that you wanted to read, what part of the citation would you use to see if the library owned the publication?
Oberman, Kieran. "Poverty and Immigration Policy." *American Political Science Review* 109, no. 2 (2015): 239–51.

> Search for the article title, "Poverty and Immigration Policy"
> *Check the Journal List on the library website for the periodical, American Political Science Review*
> Search the author Oberman, K.
> Search the subjects "immigration" AND "poverty"

*Correct answers are in italics. For questions with more than one correct response, partial credit is awarded for each correct response and deducted for each incorrect response.

---

vary each time the test is taken. Overall, JOLT provides especially good coverage of LO1 and LO2. JOLT therefore complements the assessment of students' written work, which focuses on learning objectives 3, 4, and 5.

The library has administered the first-year JOLT test every year since 2014–2015. The same test is a better assessment instrument when taken in the final year, however. As a senior-year test, it gauges long-term retention of IL knowledge, demonstrates the capabilities of graduating seniors, disentangles assessment from the process by which students demonstrate their first-year knowledge of IL fundamentals, and uses a sampling procedure that does not require faculty participation (unlike assessment of seniors' written work). When taken by seniors, JOLT is known as BOLT.

In April 2019, all graduating seniors were sent an e-mail message inviting them to take the test. Participation was not associated with enrollment in any course. To encourage students to take the test seriously, the library offered prizes to a random sample of test-takers, a random sample of those who scored 80 percent or higher, and a random sample of those who scored 90 percent or higher. Everyone who took the test had some chance of winning a prize, but the odds were higher for those who did well. Getting students to take the test seriously is important, since even a small number of unmotivated examinees can bias the results.[42] As Table 3 shows, 105 seniors took the BOLT test in spring 2019.

> **Getting students to take the test seriously is important, since even a small number of unmotivated examinees can bias the results.**

Reporting of Results

Each year, the library reports the number of students who took the first-year JOLT one or more times, the number of first attempts, and the number of repeat attempts. In some years, it also reported the retake rate as well as the number of students who took the test two, three, or more than three times.

Because each student must retake the JOLT test until he or she passes it, there are at least three sets of scores that could be included: all scores (from all attempts); the first-attempt scores; and the final-attempt scores, which are presumably the highest, since a student who passes the test need not take it again. Moodle provides summary data for all three sets of scores. The library's reports have focused primarily on the final-attempt scores, since the library is most interested in whether students have learned the material by the end of the first year—not whether they know it when they first arrive on campus. However, the library also reports the first-attempt scores to evaluate whether its first-year IL program produces immediate improvements in performance. The results include the average overall percentage score (called the *facility index* in Moodle) as well as the scores for the JOLT questions associated with each of the five learning objectives (see Table 9). Although Table 9 shows the JOLT results for several years combined, the library more often reports data from individual years to highlight changes over time. The same table format is used for BOLT, but with BOLT there is no need to distinguish between first and final attempts.

Although the library does not present data for individual JOLT and BOLT questions in its annual reports, it does look closely at those results. For instance, students tend to do poorly on the questions that involve both LO2 (searching and retrieving documents) and LO5 (citing sources and using citations). Most of the students can construct citations, but many find it difficult to understand and use the citations they find—to retrieve articles based on incomplete citations, for instance. In response, the library has increased its coverage of that topic.

> **Most of the students can construct citations, but many find it difficult to understand and use the citations they find.**

Third Assessment Method: Students' Evaluations of Library Instruction

Assessment Activities

Librarians who teach IL in the first-year composition course are expected to distribute student evaluation forms in most of their classes. The forms have two purposes: to identify potential areas of improvement for individual librarians and to solicit students' opinions of how well particular learning objectives have been achieved. These perception-based student evaluations are secondary to the other two components of the assessment program. As indirect measures of cognitive outcomes, they are subject to the limitations and biases associated with self-reports. Moreover, student evaluations are linked to just one aspect of the instruction program and cannot account for students' long-term retention of IL knowledge. Nonetheless, the library does interpret the student evaluation data to discover trends or problems that might not be apparent from students' written work and JOLT or BOLT scores. Although students are not required to complete the evaluation forms, nearly all of them do.

The evaluation form asks students to complete five statements:

Table 9.

Example table: First-year JOLT* scores, overall and for each of the
five learning objectives (LOs), spring 2015 through 2018–2019

Statistic	Overall	LO1	LO2	LO3	LO4	LO5
First attempt:						
Average score (%)	79.0	77.9	76.7	84.3	76.0	74.0
Standard deviation†	9.0	14.8	12.2	12.2	19.4	15.6
% scoring ≥ 60%	97.2	84.9	90.8	95.7	80.7	81.8
% scoring ≥ 70%	85.9	74.8	70.4	85.4	66.8	63.8
% scoring ≥ 80%	47.6	49.7	40.1	59.0	41.2	38.8
% scoring ≥ 90%	10.5	18.7	13.4	31.4	29.4	12.1
Final attempt:						
Average score (%)	81.3	80.1	79.2	86.3	78.4	76.9
Standard deviation	6.9	13.5	10.3	10.6	17.5	13.7
% scoring ≥ 60%	99.9	89.3	96.6	98.4	85.1	88.5
% scoring ≥ 70%	99.2	80.8	79.1	90.8	71.2	71.4
% scoring ≥ 80%	54.2	55.0	45.3	65.5	44.5	44.0
% scoring ≥ 90%	12.0	21.4	15.4	35.5	32.3	14.0

*JOLT (Jasper Online Information Literacy Test) is Manhattan College's first-year test of IL
proficiency.
†The standard deviation represents the dispersion or variability of the scores.

> With what I've learned in this instruction session, I now feel confident that I can . . . (1)
> understand the different kinds of information available for my research assignment,
> (2) find appropriate resources for my assignment, (3) evaluate the appropriateness
> of resources for my assignment, (4) appropriately cite the information I use in my
> assignment, and (5) use a citation to find a specific book, article or website.

The first three statements are keyed to the first three learning objectives, and the last
two match LO5 (citing sources and using citations). Learning objective 4 (integrating
information into academic work) is not assessed through the student evaluations. For
each statement, the student selects a response of 1 (strongly disagree), 2 (disagree), 4
(agree), or 5 (strongly agree).

The form also includes several open-ended items that encourage students to reflect
on the IL session, asking them to complete sentences that begin "I'm especially glad you
covered . . . "; "I wish you could have included more information about . . . "; "I'm still
confused about . . . "; and "I especially liked your presentation/discussion of . . . " Each
year, these items are evaluated through an informal content analysis, with special at-

tention to responses mentioned by more than a few students or that suggest difficulties or strengths not captured by other methods.

Reporting of Results

For each of the five scaled-response items, the library's annual report presents the average value; the standard deviation, which represents the dispersion or variability of the scores; the number of responses; and the percentage of students giving responses of 4 (agree) or 5 (strongly agree). A significant problem, from an assessment perspective, is the high number of 4 or 5 responses—generally more than 95 percent, and sometimes as high as 100 percent. The average scores for all questions are almost uniformly in the 4.3 to 4.5 range. The high scores and low standard deviations suggest that these items are not useful for distinguishing between effective and less effective instruction, even from the students' perspective.

The open-ended comments are more useful, however, since they indicate where students' comprehension is less than ideal and suggest topics for which more intensive or more advanced instruction would be helpful. For instance, students' comments have reinforced an idea seen earlier in the JOLT results—that they have more difficulty using citations than generating them. The reports do not usually include tables showing the content analysis results, but the library does rely on those results when determining which topics to mention in the text of each year's report.

Using the Assessment Results

Main Findings

The main findings are remarkably consistent. Most were revealed through multiple assessment methods over the course of the five-year period and can be seen in both the first-year and senior-year results. The situation at Manhattan College may not be directly relevant to other institutions, but the following types of findings emerged from the assessment process.

In general, Manhattan College students appear to have no major difficulties with LO1 (identifying information needs) and LO3 (evaluating and selecting documents). This may reflect the library's longstanding emphasis on those two learning objectives within the English composition IL sessions. Students do well in terms of both 3.1 (relevance of sources to the student's needs) and 3.2 (quality of sources), and seniors attain higher scores than first-year students. This suggests that students improve their LO3 capabilities over the course of their undergraduate years. For LO2 (searching and retrieving documents), students' performance is mostly satisfactory. However, students tend to do poorly on tasks that involve the combination of LO2 and LO5 (citing sources and using citations).

LO4 (integrating information into academic work) is the one learning objective for which the students' performance is consistently unsatisfactory. The problem is apparent for both first-year students and seniors, and it can be seen in both 4.1 (degree to which the student has incorporated information resources when it is necessary or appropriate to do so) and 4.2 (task-specific relevance and use of sources). Although most

students understand when it is necessary to cite the information they have used, fewer understand when it is necessary or appropriate to use published information in support of their narratives or arguments. (See the Appendix, item 4.1). When discussing students' written work, librarians noted the number of times they would have written "evidence?" The problem can be seen not just for assertions that might be contested ("In Eastern Europe, stricter gun control laws have led to a reduction in the rate of violent crime") but also for simple factual statements ("Super Bowl viewership first exceeded 100 million in 2010"). Over the past three years, nearly half of all seniors have received unsatisfactory scores (1 or 2) for item 4.1 (degree to which the student has incorporated information resources when it is necessary or appropriate to do so).

> Although most students understand when it is necessary to cite the information they have used, fewer understand when it is necessary or appropriate to use published information in support of their narratives or arguments.

Manhattan College students also encounter difficulties with 4.2 (task-specific relevance and use of sources). Their reasons for citing a source are not always apparent. Some students tend to cite works in passing, without any meaningful information about the cited authors' arguments or findings; to include works that are topically relevant but do not support or clarify the student's presentation or argument; to omit the descriptive text that might provide a context for the cited author's findings; and to describe each cited work in a single, self-contained paragraph without any attempt to address the differences between works, the common themes that emerge within multiple works, or the ways in which any particular work relates to the others. Fortunately, seniors are less likely than first-year students to make these kinds of errors. Requests for further instruction on topics such as "reading articles for relevant content" were encountered repeatedly in students' evaluations, so at least some students recognize the need to improve their capabilities regarding LO4 (integrating information into academic work).

Manhattan College students do reasonably well with LO5 (citing sources and using citations). As noted earlier, however, they are better at constructing citations than at interpreting them or using them to retrieve relevant works. The library's assessments reveal a high standard deviation for the LO5 scores, which suggests that some students learn the material far better than others. While many students' papers included complete, correctly formatted citations, a significant minority of students failed to provide the bibliographic elements that readers would need to evaluate the adequacy of their sources. Some students cited websites without giving the URLs, for example.

> While many students' papers included complete, correctly formatted citations, a significant minority of students failed to provide the bibliographic elements that readers would need to evaluate the adequacy of their sources.

Nearly all the assessment results can be interpreted regarding the five learning objectives. Two unexpected findings did emerge, however. One is the realization that, in recent years, first-year students who fail the JOLT on their first attempt sometimes do worse on their second attempt before eventually passing the test. This suggests that they do not follow the advice to consult with a librarian or take the online tutorial before attempting the examination again. Second, some students mentioned "organization of books in the library" when asked what ought to be covered more fully in their first-year IL instruction. Students appear to want more help locating print books (that is, call numbers), something the library had not previously felt the need to cover.

"Closing the Loop"

The goal of assessment is not simply to gauge students' capabilities but also to improve their performance. ("Closing the loop," in this context, refers to the process of using assessment data to improve teaching, research, or service.) Demonstrated improvement, rather than satisfactory performance, may be the preferred result for many accreditors as well.[43]

At Manhattan College, five years of assessment have led to substantial changes in the IL instruction program. The process first drew attention to the difficulty of covering the five learning objectives, even at a basic level, in a single session of English composition (50–75 minutes). In response, the library worked with the writing program to increase the number of faculty who devote two or more class sessions to IL instruction. The current average is 1.4 sessions, and the library hopes to increase that number in the coming years. The additional time will allow librarians to cover the learning objectives more completely and to provide more opportunities for discussion, activities, games, group work, and other types of active learning. Time constraints have sometimes led librarians to use a lecture format even when other methods might have been more effective at engaging students and helping them learn.

The assessment process has also led the library to evaluate its online research guides—to improve existing guides, create new ones, and remove those that do not meet current standards. Although this change was not linked to any one assessment result, it was embedded in the broader process of evaluating the library's

> **Time constraints have sometimes led librarians to use a lecture format even when other methods might have been more effective at engaging students and helping them learn.**

instructional efforts and ensuring that they support all five learning objectives. The library probably would have reviewed its research guides with or without the assessment results, but the assessment program provided a structure that made the review process more deliberate and more effective. In addition, the library has (1) met repeatedly with the English faculty to help them understand the context and goals of the IL program, (2) updated and revised the JOLT tutorial to bring it into closer alignment with the library's learning objectives, (3) modified the student evaluation forms and peer observation forms to address the five learning objectives more directly, (4) established a series of workshops

in which librarians and other instructional personnel discuss methods and issues related to teaching and IL, and (5) renewed its efforts to ensure that students who fail JOLT will meet with a librarian or take the JOLT tutorial before attempting the test a second time.

Other improvements were made in response to specific assessment results. The library has tried to address students' difficulties integrating information into their academic work (LO4) and using citations to find and access relevant books and articles (the intersection of LO2, searching and retrieving documents, and LO5, citing sources and using citations). It has expanded its in-class coverage of both topics and added new citation-related content to many of the library research guides. Coverage of LO4 will likely increase further with the additional time available in composition classes. LO4 instruction cannot be undertaken by the librarians alone, however, since it involves methods and expectations that are tied to the content of each course. In addition, students will likely require repeated exposure to best practices as well as repeated reminders—instructors' comments on their papers and assignments, for instance.[44] The librarians have begun to work more closely with faculty, including those teaching courses other than composition, to ensure that the rationale and expectations for LO4 are understood and supported.

Additional changes have been planned but not yet undertaken. In the coming years, the library will (1) increase the number of stand-alone workshops devoted to LO4 and LO5;[45] (2) create online research guides on citing and integrating information into academic work; (3) develop short videos and other point-of-need instructional materials;[46] (4) build and maintain a shared collection of teaching materials, such as handouts and activity guides, that support specific learning objectives; and (5) investigate opportunities for IL instruction beyond the first year so that the concepts taught in English composition can be reinforced, and so that higher level material can be taught more effectively in preparation for students' junior- and senior-year papers. The last item is the most ambitious, especially for a library that has always concentrated on first-year instruction.

Because the library's IL learning objectives guide both instruction and assessment, these instructional changes are expected to improve students' performance over time. So far, however, only small improvements can be demonstrated. For instance, seniors' written work shows some progress over the past few years for LO5 (citing sources and using citations). This finding was surprising. Although the librarians have increased their coverage of LO5 in first-year composition, they did not expect to see improvements in seniors' performance until recent first-year students reached their senior year. The addition of new citation content to the online research guides may have already had an impact on students at all levels, and the emphasis on citations in first-year instruction may have been informally integrated into work with upper-division students.

Further Assessment Strategies

As mentioned earlier, it has been difficult to get representative samples of seniors' written work. In no single year did the library receive papers from all six of the Manhattan College schools. Because different schools have participated each year, however, it is possible to obtain a more representative sample by pooling (combining) three years' data. This is justifiable due to the low variation in seniors' performance over time. In pooling the data, the library regards 2015–2016 through 2017–2018 as a single pre-improvement

period. After two years of instructional improvements ("closing the loop"), it will resume its assessment program in the post-improvement period (2021–2022 and thereafter) by evaluating seniors' written work, interpreting students' JOLT and BOLT scores, and examining their evaluations of in-class IL instruction.

The library's assessment efforts have benefited from the designation of IL as a college-wide core competency. That designation has allowed it to frame its assessment program as something beneficial to, and required by, the college as a whole. Information literacy was the first of seven core competencies assessed at Manhattan College, and the program has since served as a model for the assessment of students' written and oral communication skills. Gaining the cooperation of the faculty has been difficult, however, especially regarding the evaluation of senior papers and the use of class time for IL instruction. Although the college administration fully supports the library's IL initiatives, individual faculty sometimes question the need to give students

> **Although the college administration fully supports the library's IL initiatives, individual faculty sometimes question the need to give students more than a cursory overview of the library website.**

more than a cursory overview of the library website. Likewise, more than one composition instructor has claimed that they already teach information literacy, despite the minimal overlap between the IL and communication core competencies.[47]

Overall, the library's IL assessment program has been effective regarding the first six items presented in the Introduction. However, it has not yet evaluated the effectiveness of instructional methods and materials, systematically identified the students in greatest need of further instruction, or examined the relationships between IL competencies and broader educational outcomes. The results have helped determine what needs to be taught more intensively, but not how to teach it. Likewise, the library targets the students in greatest need of assistance only to the extent that they choose to make use of its research and reference services. Finally, the library has not yet attempted to link its assessment results to students' academic records. Such an approach holds great promise, however, since it would allow the library to evaluate the relationships between IL expertise and indicators such as course grades, retention rates, and career outcomes.[48]

Although the Manhattan College library cannot claim that its assessment program is more effective than others, it has proved valuable as a means of identifying the areas in which additional instruction is needed. Moreover, these methods can be implemented without great difficulty at a wide range of colleges and universities. Many components of the program build on activities that are routinely undertaken for other purposes, such as IL testing and student evaluations. The assessment methods are also scalable. As noted in Table 7, the size of the undergraduate population has no bearing on the number of survey responses or term papers needed for a reliable analysis.

Acknowledgments

William H. Walters, a member of the Editorial Board of portal: Libraries and the Academy*, is executive director of the library at Manhattan College in Riverdale, New York; he may be reached by e-mail at: william.walters@manhattan.edu.*

Sarah E. Sheehan is assistant director of the library for reference and instruction at Manhattan College in Riverdale, New York; she may be reached by e-mail at: sarah.sheehan@manhattan.edu.

Amy E. Handfield is assistant director of the library for access services at Manhattan College in Riverdale, New York; she may be reached by e-mail at: amy.handfield@manhattan.edu.

Bernadette M. López-Fitzsimmons is an associate librarian at Manhattan College in Riverdale, New York; she may be reached by e-mail at: bernadette.lopez@manhattan.edu.

Susanne Markgren is assistant director of the library for technical services at Manhattan College in Riverdale, New York; she may be reached by e-mail at: susanne.markgren@manhattan.edu.

Laurin Paradise is a reference and instruction librarian at Manhattan College in Riverdale, New York; she may be reached by e-mail at: laurin.paradise@manhattan.edu.

Appendix

Assessment Rubric Used to Evaluate Students' Written Work

The six assessment criteria included in the rubric are keyed to learning objectives 3, 4, and 5 (Table 2). Although the criteria and levels of performance are shown here as bullet points, the assessment rubric is usually presented in grid (table) format, with the six criteria as row headings and the four levels of performance as column headings. The complete rubric also includes the text of the learning objectives.

The four performance levels are

1. Below minimum standard—major problems
2. Below minimum standard—some problems
3. Meets minimum standard—essentially proficient
4. Exceeds minimum standard—fully proficient.

3.1 Relevance of sources to the student's needs—particular items chosen

1. Does not meet the requirements for a score of 2.
2. At least 50 percent of the sources address the student's specific research problem, argument, or question.
3. At least 75 percent of the sources address the student's specific research problem, argument, or question and at least 50 percent are functionally relevant.
4. Nearly all the sources address the student's specific research problem, argument, or question, and nearly all are functionally relevant.

3.2 Quality of sources—particular items chosen

1. Does not meet the requirements for a score of 2.
2. At least 50 percent of the sources are satisfactory in terms of currency, authority, and bias.
3. At least 75 percent of the sources are satisfactory in terms of currency, authority, and bias. If biased sources are used, their use is appropriate and intentional, and if the instructor has specified a minimum number or type of sources, that requirement has been met.
4. Nearly all the sources are satisfactory in terms of currency, authority, and bias. If biased sources are used, their use is appropriate and intentional, and if the instructor has specified a minimum number or type of sources, that requirement has been met.

4.1 Degree to which student has incorporated information resources when it is necessary or appropriate to do so

1. Does not meet the requirements for a score of 2.
2. There are many instances in which the student has failed to incorporate information that is necessary or obviously appropriate to his or her objective (e.g., fails to cite evidence in support of an assertion).
3. There are only a few instances in which the student has failed to incorporate information that is necessary or obviously appropriate to his or her objective (e.g., fails to cite evidence in support of an assertion).
4. There are almost no instances in which the student has failed to incorporate information that is necessary or obviously appropriate to his or her objective (e.g., fails to cite evidence in support of an assertion), and information sources are cited in nearly all the places where they might be expected by an educated reader.

4.2 Task-specific relevance and use of sources

1. Does not meet the requirements for a score of 2.
2. In at least 50 percent of the instances in which information is used/cited, the underlying rationale for its use is apparent.
3. In at least 50 percent of the instances in which information is used/cited, the underlying rationale for its use is apparent and the student introduces or explains the cited information in most cases, avoiding simple listings of studies and their findings.
4. The underlying rationale for the use of information is always clear, the student introduces or explains the cited information well, and where appropriate, the student has compiled the information in ways that serve his or her goals (e.g., in a table that show the key findings of multiple studies). (Require high standards for a score of 4.)

5.1 Attribution and identification of cited works

1. Does not meet the requirements for a score of 2.
2. At least 50 percent of the sources cited or mentioned in the text appear (in some form) in the Works Cited section, at least 50 percent of the sources listed in the Works Cited section are cited or mentioned in the text, and at least 50 percent of the citations include enough information for the identification of the work by the assessor.
3. At least 75 percent of the sources cited or mentioned in the text appear (in some form) in the Works Cited section, at least 75 percent of the sources listed in the Works Cited section are cited or mentioned in the text, and at least 75 percent of the citations include enough information for the identification of the work by the assessor.
4. Nearly all the sources cited or mentioned in the text appear (in some form) in the Works Cited section, nearly all the sources listed in the Works Cited section are cited or mentioned in the text, and nearly all the citations include enough information for the identification of the work by the assessor.

5.2 Completeness and format of citations

1. Does not meet the requirements for a score of 2.
2. At least 50 percent of the in-text citations are correctly formatted, and at least 50 percent of the entries in the Works Cited section include the bibliographic information specified by the citation style.
3. At least 75 percent of the in-text citations are correctly formatted, at least 75 percent of the entries in the Works Cited section include the bibliographic information specified by the citation style, and stylistic elements (indentation, punctuation, capitalization, alphabetization, order of elements) are often correct or nearly correct.

4. Nearly all the in-text citations are correctly formatted, nearly all the entries in the Works Cited section include the bibliographic information specified by the citation style, and stylistic elements (indentation, punctuation, capitalization, alphabetization, order of elements) are nearly always correct.

Notes

1. This conclusion is based on title words, which provide a more reliable count than subject headings. The ERIC descriptors *educational assessment* and *evaluation* (used for *assessment*) have not been applied consistently over time.

2. Corey M. Johnson, Carol M. Anelli, Betty J. Galbraith, and Kimberly A. Green, "Information Literacy Instruction and Assessment in an Honors College Science Fundamentals Course," *College & Research Libraries* 72, 6 (2011): 533–47; Thomas P. Mackey and Trudi E. Jacobson, "Developing an Integrated Strategy for Information Literacy Assessment in General Education," *Journal of General Education* 56, 2 (2007): 93–104; Paula McMillen and Anne-Marie Deitering, "Complex Questions, Evolving Answers: Creating a Multidimensional Assessment Strategy to Build Support for the 'Teaching Library," *Public Services Quarterly* 3, 1–2 (2007): 57–82; Kornelia Tancheva, Camille Andrews, and Gail Steinhart, "Library Instruction Assessment in Academic Libraries," *Public Services Quarterly* 3, 1–2 (2007): 29–56. For brief descriptions of nearly 200 other IL assessment strategies, see Association of College and Research Libraries (ACRL), "Assessment in Action," 2019, https://apply.ala.org/aia/public.

3. A more detailed description of the assessment cycle is provided by Megan Oakleaf, "The Information Literacy Assessment Cycle: A Guide for Increasing Student Learning and Improving Librarian Instructional Skills," *Journal of Documentation* 65, 4 (2009): 539–60.

4. The best two books on academic program assessment, intended for practitioners but thoroughly grounded in the social science literature, are Mary J. Allen, *Assessing General Education Programs* (Bolton, MA: Anker, 2006); and Trudy W. Banta and Catherine A. Palomba, *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education*, 2nd ed. (San Francisco, CA: Jossey-Bass, 2015). Also see Peter Hernon, Robert E. Dugan, and Candy Schwartz, eds., *Higher Education Outcomes Assessment for the Twenty-First Century* (Santa Barbara, CA: Libraries Unlimited, 2013); Richard J. Shavelson, *Measuring College Learning Responsibly: Accountability in a New Era* (Stanford, CA: Stanford University Press, 2009); Linda Suskie, *Assessing Student Learning: A Common Sense Guide*, 2nd ed. (San Francisco, CA: Jossey-Bass, 2018).

5. ACRL, "Academic Library Statistics," 2019, http://www.ala.org/acrl/publications/trends; National Center for Education Statistics, "IPEDS: Integrated Postsecondary Education Data System," 2019, https://nces.ed.gov/ipeds.

6. Alexander W. Astin and Anthony Lising Antonio, *Assessment for Excellence: The Philosophy and Practice of Assessment and Evaluation in Higher Education*, 2nd ed. (Lanham, MD: Rowman & Littlefield, 2012).

7. Leslie Owen Wilson, "Three Domains of Learning—Cognitive, Affective, Psychomotor," in *The Second Principle: The Work of Leslie Owen Wilson*, 2019, https://thesecondprinciple.com/instructional-design/threedomainsoflearning/.

8. Southern Methodist University, "Direct and Indirect Measures," 2019, https://www.smu.edu/Provost/assessment/Measures.

9. ACRL, "Assessment in Action." For additional case studies, see Marwin Britto and Kirsten Kinsley, eds., *Academic Libraries and the Academy: Strategies and Approaches to Demonstrate Your Value, Impact, and Return on Investment*, 2 vols. (Chicago: American Library Association [ALA], 2018).

10. Murray State University in Murray, Kentucky, the University of Massachusetts–Dartmouth, and the University of West Georgia in Carrollton were excluded from the summary data because their assessment programs do not have an IL instruction component.

11. A 2006 study of Australian postsecondary vocational libraries reported far less emphasis on the direct assessment of written work; see Jenny Fafeita, "The Current Status of Teaching and Fostering Information Literacy in TAFE [technical and further education]," *Australian Academic & Research Libraries* 37, 2 (2006): 136–61.

12. Washington State University in Pullman, the State University of New York at Albany (SUNY Albany), Oregon State University in Corvallis, and Cornell University in Ithaca, New York, have also developed assessment programs that rely on multiple methods, including both direct and indirect assessments of cognitive and affective outcomes; see Johnson, Anelli, Galbraith, and Green, "Information Literacy Instruction and Assessment in an Honors College Science Fundamentals Course"; Mackey and Jacobson, "Developing an Integrated Strategy for Information Literacy Assessment in General Education"; McMillen and Deitering, "Complex Questions, Evolving Answers"; Tancheva, Andrews, and Steinhart, "Library Instruction Assessment in Academic Libraries."

13. Middle States Commission on Higher Education (MSCHE), "Standard V: Educational Effectiveness Assessment," in *Standards for Accreditation and Requirements of Affiliation*, 13th ed. (Philadelphia: MSCHE, 2018), 10–11, http://www.msche.org/wp-content/uploads/2018/06/RevisedStandardsFINAL.pdf; Western Association of Schools and Colleges (WASC), "Standard 2: Achieving Educational Objectives through Core Functions," in *Handbook of Accreditation 2013 Revised* (Alameda, CA: WASC, 2018), https://www.wscuc.org/resources/handbook-accreditation-2013/part-ii-core-commitments-and-standards-accreditation/wasc-standards-accreditation-2013/standard-2-achieving-educational-objectives-through-core-functions.

14. Jane Hiscock, "Does Library Usage Affect Academic Performance?" *Australian Academic and Research Libraries* 17, 4 (1986): 207–14; John K. Stemmer and David M. Mahan, "Investigating the Relationship of Library Usage to Student Outcomes," *College & Research Libraries* 77, 3 (2016): 359–75; Krista M. Soria, Jan Fransen, and Shane Nackerud, "Library Use and Undergraduate Student Outcomes: New Evidence for Students' Retention and Academic Success," *portal: Libraries and the Academy* 13, 2 (2013): 147–64; Shun Han Rebekah Wong and Dianne Cmor, "Measuring Association between Library Instruction and Graduation GPA," *College & Research Libraries* 72, 5 (2011): 464–73; Shun Han Rebekah Wong and T. D. Webb, "Uncovering Meaningful Correlation between Student Academic Performance and Library Material Usage," *College & Research Libraries* 72, 4 (2011): 361–70.

15. This distinction has been made clearly by Szarina Abdullah, "Measuring the Outcomes of Information Literacy: Perception vs. Evidence-Based Data," *International Information & Library Review* 42, 2 (2010): 98–104.

16. Association of Research Libraries, "LibQUAL+™: Charting Library Service Quality," 2019, https://www.libqual.org/home.

17. Maria Pinto, "Design of the IL-HUMASS [Information Literacy Humanities and Social Sciences] Survey on Information Literacy in Higher Education: A Self-Assessment Approach," *Journal of Information Science* 36, 1 (2010) 86–103; María Pinto, "An Approach to the Internal Facet of Information Literacy Using the IL-HUMASS Survey," *Journal of Academic Librarianship* 37, 2 (2011): 145–54.

18. Amy E. Mark and Polly D. Boruff-Jones, "Information Literacy and Student Engagement: What the National Survey of Student Engagement Reveals about Your Campus," *College & Research Libraries* 64, 6 (2003): 480–93.

19. See, for example, Brian Detlor, Heidi Julien, Rebekah Willson, Alexander Serenko, and Maegen Lavallee, "Learning Outcomes of Information Literacy Instruction at Business Schools," *Journal of the American Society for Information Science and Technology* 62, 3 (2011): 572–85; Donald Gilstrap and Jason Dupree, "Assessing Learning, Critical Reflection, and Quality Educational Outcomes: The Critical Incident Questionnaire," *College & Research Libraries* 69, 5 (2008): 407–26; Gillian Gremmels and Karen Lehmann, "Assessment of Student Learning from Reference Service," *College & Research Libraries* 68, 6 (2007): 488–501; Kate Zoellner, Sue Samson, and Samantha Hines, "Continuing Assessment of Library

Instruction to Undergraduates: A General Education Course Survey Research Project," *College & Research Libraries* 69, 4 (2008): 370–83.

20. Peter Hernon and Ellen Altman, *Assessing Service Quality: Satisfying the Expectations of Library Customers*, 2nd ed. (Chicago: ALA, 2010); Bruce Thompson, Colleen Cook, and Russel L. Thompson, "Reliability and Structure of LibQUAL+ Scores: Measuring Perceived Donald L. Gilstrap and Jason Dupree, "Assessing Learning, Critical Reflection, and Quality Educational Outcomes: The Critical Incident Questionnaire," *College & Research Libraries* 69, 5 (2008): 407–26.

21. Abdullah, "Measuring the Outcomes"; Katherine Schilling and Rachel Applegate, "Evaluating Library Instruction: Measures for Assessing Educational Quality and Impact," in *Proceedings of the ACRL Thirteenth National Conference* (Chicago: ALA, 2007), 206–14; Randall Schroeder and Kimberly Babcock Mashek, "Building a Case for the Teaching Library: Using a Culture of Assessment to Reassure Converted Campus Partners while Persuading the Reluctant," *Public Services Quarterly* 3, 1–2 (2007): 83–110; William H. Walters, "Expertise and Evidence in the Assessment of Library Service Quality," *Performance Measurement and Metrics* 4, 3 (2003): 98–102; William H. Walters, "Beyond Use Statistics: Recall, Precision, and Relevance in the Assessment and Management of Academic Libraries," *Journal of Librarianship and Information Science* 48, 4 (2016): 340–52.

22. Anne Boring, "Gender Biases in Student Evaluations of Teaching," *Journal of Public Economics* 145 (2017): 27–41; Friederike Mengel, Jan Sauermann, and Ulf Zölitz, "Gender Bias in Teaching Evaluations," *Journal of the European Economic Association* 17, 2 (2019): 535–66; Kristina M.W. Mitchell and Jonathan Martin, "Gender Bias in Student Evaluations," *PS: Political Science & Politics* 51, 3 (2018): 648–52.

23. Walters, "Beyond Use Statistics."

24. William J. Goode, "Community within a Community: The Professions," *American Sociological Review* 22, 2 (1957): 194–200.

25. Rachel Applegate, "Models of User Satisfaction: Understanding False Positives," *RQ* 32, 4 (1993): 525–39; Patricia Davitt Maughan, "Assessing Information Literacy among Undergraduates: A Discussion of the Literature and the University of California-Berkeley Assessment Experience," *College & Research Libraries* 62, 1 (2001): 71–85; T. Scott Plutchak, "On the Satisfied and Inept End User," *Medical Reference Services Quarterly* 8, 1 (1989): 45–48; Katherine Schilling and Rachel Applegate, "Best Methods for Evaluating Educational Impact: A Comparison of the Efficacy of Commonly Used Measures of Library Instruction," *Journal of the Medical Library Association* 100, 4 (2012): 258–69; Robert S. Weisskirch and Janie B. Silveria, "The Effectiveness of Project-Specific Information Competence Instruction," *Research Strategies* 20, 4 (2005): 370–78.

26. Schilling and Applegate, "Evaluating Library Instruction."

27. Peggy Brady-Amoon and Jairo N. Fuertes, "Self-Efficacy, Self-Rated Abilities, Adjustment, and Academic Performance," *Journal of Counseling & Development* 89, 4 (2011): 431–38; Angelo S. DeNisi and James B. Shaw, "Investigation of the Uses of Self-Reports of Abilities," *Journal of Applied Psychology* 62, 5 (1977): 641–44; Toni Honicke and Jaclyn Broadbent, "The Influence of Academic Self-Efficacy on Academic Performance: A Systematic Review," *Educational Research Review* 17 (2016): 63–84.

28. Robert Schroeder and Ellysa Stern Cahoy, "Valuing Information Literacy: Affective Learning and the ACRL Standards," *portal: Libraries and the Academy* 10, 2 (2010): 127–46. Cognitive and affective assessments can also be used together to identify cases in which abilities and attitudes diverge. Arguably, the worst outcome for an IL instruction program is that graduates will have poor IL skills yet still feel confident in their abilities.

29. Lynn Cameron, Steven L. Wise, and Susan M. Lottridge, "The Development and Validation of the Information Literacy Test," *College & Research Libraries* 68, 3 (2007): 229–36; Carrick Enterprises, "Project SAILS [Standardized Assessment of Information Literacy Skills]: The Individual Scores Test," 2019, https://www.projectsails.org/site/the-individual-scores-test/; Madison Assessment, "Information Literacy Test," 2019, https://www.madisonassessment.com/assessment-testing/information-literacy-test.

30. Penny M. Beile O'Neil, "Development and Validation of the Beile Test of Information Literacy for Education (B-TILED)," PhD diss., University of Central Florida, 2005, http://purl.fcla.edu/fcla/etd/CFE0000749; Lana Ivanitskaya, Irene O'Boyle, and Anne Marie Casey, "Health Information Literacy and Competencies of Information Age Students: Results from the Interactive Online Research Readiness Self-Assessment (RRSA)," *Journal of Medical Internet Research* 8, 2 (2006): e6.

31. Jacalyn E. Bryan and Elana Karshmer, "Assessment in the One-Shot Session: Using Pre- and Post-Tests to Measure Innovative Instructional Strategies among First-Year Students," *College & Research Libraries* 74, 6 (2013): 574–86; Elizabeth Choinski and Michelle Emanuel, "The One-Minute Paper and the One-Hour Class: Outcomes Assessment for One-Shot Library Instruction," *Reference Services Review* 34, 1 (2006): 148–55; Michelle Kathleen Dunaway and Michael Teague Orblych, "Formative Assessment: Transforming Information Literacy Instruction," *Reference Services Review* 39, 1 (2011): 24–41; Jon R. Hufford, "What Are They Learning? Pre- and Post-Assessment Surveys for LIBR 1100, Introduction to Library Research," *College & Research Libraries* 71, 2 (2010): 139–58; Schroeder and Mashek, "Building a Case for the Teaching Library."

32. For a comparison of tests and other cognitive assessment methods, see Megan Oakleaf, "Dangers and Opportunities: A Conceptual Map of Information Literacy Assessment Approaches," *portal: Libraries and the Academy* 8, 3 (2008): 233–53.

33. Association of American Colleges & Universities (AAC&U), "Information Literacy VALUE [Valid Assessment of Learning in Undergraduate Education] Rubric," 2013, https://www.aacu.org/value/rubrics/information-literacy; Jessica Alverson, Jennifer Schwartz, and Sue Shultz, "Authentic Assessment of Student Learning in an Online Class: Implications for Embedded Practice," *College & Research Libraries* 80, 1 (2019): 32–43; Karen R. Diller and Sue F. Phelps, "Learning Outcomes, Portfolios, and Rubrics, Oh My! Authentic Assessment of an Information Literacy Program," *portal: Libraries and the Academy* 8, 1 (2008): 75–89; Debra Anne Hoffmann and Kristen LaBonte, "Meeting Information Literacy Outcomes: Partnering with Faculty to Create Effective Information Literacy Assessment," *Journal of Information Literacy* 6, 2 (2012): 70–85; Oakleaf, "The Information Literacy Assessment Cycle"; Davida Scharf, Norbert Elliot, Heather A. Huey, Vladimir Briller, and Kamal Joshi, "Direct Assessment of Information Literacy Using Writing Portfolios," *Journal of Academic Librarianship* 33, 4 (2007): 462–77. The RAILS (Rubric Assessment of Information Literacy Skills) site may be especially useful to librarians or faculty seeking to develop and validate institution-specific IL rubrics; see Megan Oakleaf, "RAILS: Rubric Assessment of Information Literacy Skills," 2019, http://railsontrack.info/; Megan Oakleaf, "Staying on Track with Rubric Assessment: Five Institutions Investigate Information Literacy Learning," *Peer Review* 13, 4 (2011): 18–21.

34. See, for example, Walters, "Beyond Use Statistics"; William H. Walters and Susanne Markgren, "Do Faculty Journal Selections Correspond to Objective Indicators of Citation Impact? Results for 20 Academic Departments at Manhattan College," *Scientometrics* 118, 1 (2019): 321–37.

35. AAC&U, "Information Literacy VALUE Rubric"; ACRL, "Information Literacy Competency Standards for Higher Education," 2000, http://hdl.handle.net/11213/7668; ACRL, "Characteristics of Programs of Information Literacy That Illustrate Best Practices: A Guideline," *College & Research Libraries News* 73, 6 (2012): 355–59.

36. ACRL, "Framework for Information Literacy for Higher Education," 2016, http://www.ala.org/acrl/standards/ilframework.

37. Debra Cox Rollins, Jessica Hutchings, Melissa Ursula, Dawn Goldsmith, and Anthony J. Fonseca, "Are We There Yet? The Difficult Road to Re-Create Information Literacy," *portal: Libraries and the Academy* 9, 4 (2009): 453–73.

38. The Manhattan College library's tutorial is based on the Texas Information Literacy Tutorial (TILT) and the Wichita State University EMPOWER Tutorial. See Wichita State University, "EMPOWER," 2008, http://library.wichita.edu/empower/.

39. MSCHE, "Standard V"; WASC, "Standard 2."

40. To compute Cronbach's alpha, the team used the Scale—Reliability Analysis procedure in SPSS, treating each paper as a case and each rater's score for a single criterion as a variable. There were therefore 18 variables: 3 raters × 6 criteria. The team undertook six separate analyses, one for each criterion.

41. The guidelines presented in Table 7 are explained more fully in many textbooks and guides. See, for example, Gary T. Henry, *Practical Sampling* (Thousand Oaks, CA: SAGE, 1990); Graham Kalton, *Introduction to Survey Sampling* (Newbury Park, CA: SAGE, 1983); Steven K. Thompson, *Sampling*, 3rd ed. (Hoboken, NJ: Wiley, 2012).

42. Steven L. Wise, "Strategies for Managing the Problem of Unmotivated Examinees in Low-Stakes Testing Programs," *Journal of General Education* 58, 3 (2009): 152–66; Steven L. Wise and Christine E. DeMars, "Examinee Noneffort and the Validity of Program Assessment Results," *Educational Assessment* 15, 1 (2010): 27–41.

43. Accreditors face questions about the value of their work and the rationale for maintaining five regional accreditors rather than a single government agency that certifies compliance with educational standards. In this environment, accreditors may be glad to see effective instruction but even more eager to see evidence that the quality of instruction has improved due to their influence (for example, their assessment requirements and oversight activities). See, for example, Doug Lederman, "No Love, but No Alternative," *Inside Higher Ed*, September 1, 2015, https://www.insidehighered.com/news/2015/09/01/accreditation-will-change-survive.

44. Laura Saunders, "The Future of Information Literacy in Academic Libraries: A Delphi Study," *portal: Libraries and the Academy* 9, 1 (2009): 99–114.

45. See, for example, Yu-Hui Chen and Mary K. Van Ullen, "Helping International Students Succeed Academically through Research Process and Plagiarism Workshops," *College & Research Libraries* 72, 3 (2011): 209–35; Kristin Hoffmann, Fred Antwi-Nsiah, Vivian Feng, and Meagan Stanley, "Library Research Skills: A Needs Assessment for Graduate Student Workshops," *Issues in Science & Technology Librarianship* 53 (2008): article 2.

46. Although the library is open 24 hours per day, seven days per week during the academic year, the reference desk is staffed just 48 hours per week.

47. The faculty's limited understanding of IL instruction is one of the biggest challenges facing those who teach and assess information literacy. See, for example, Nancy O'Hanlon, "Information Literacy in the University Curriculum: Challenges for Outcomes Assessment," *portal: Libraries and the Academy* 7, 2 (2007): 169–89; Saunders, "The Future of Information Literacy in Academic Libraries"; Laura Saunders, "Faculty Perspectives on Information Literacy as a Student Learning Outcome," *Journal of Academic Librarianship* 38, 4 (2012): 226–36.

48. The authors know of no studies that have systematically considered the impact of measured IL capabilities on academic performance at the individual level. However, previous research has uncovered several relationships between library use (or library instruction) and academic outcomes. See Hiscock, "Does Library Usage Affect Academic Performance?"; Stemmer and Mahan, "Investigating the Relationship of Library Usage to Student Outcomes"; Soria, Fransen, and Nackerud, "Library Use and Undergraduate Student Outcomes"; Wong and Cmor, "Measuring Association between Library Instruction and Graduation GPA"; Wong and Webb, "Uncovering Meaningful Correlation between Student Academic Performance and Library Material Usage."