Dangers and Opportunities: A Conceptual Map of Information Literacy Assessment Approaches

Megan Oakleaf

abstract: The culture of assessment in higher education requires academic librarians to demonstrate the impact of information literacy instruction on student learning. As a result, many librarians seek to gain knowledge about the information literacy assessment approaches available to them. This article identifies three major assessment approaches: (1) fixed-choice tests, (2) performance assessments, and (3) rubrics. It maps the theoretical and educational assumptions on which these options are grounded and charts the dangers and opportunities of each assessment approach.

> *"When written in Chinese, the word 'crisis' is composed of two characters—*
> *one represents danger and one represents opportunity." John F. Kennedy*

Students, parents, employers, and politicians all expect colleges and universities to demonstrate excellence and commit to the pursuit of continuous improvement. In recent years, institutions of higher education have faced increased demands for evidence that they produce graduates armed with the knowledge, skills, and abilities needed to live and work in the information age. Already, academic departments and student support programs are required to assess their impact on student learning and development. Libraries have long argued that they are integral to the teaching and learning mission of higher education. Now, libraries must prove that they contribute to the production of quality graduates.

In the past, libraries have relied heavily on input, output, and process measures as key indicators of excellence. These measures are no longer considered adequate for assessing the impact of libraries' services on college students. In a climate of outcomes-based measurement, university stakeholders are less interested in traditional measurements: the count of the volumes on library shelves, the number of students checking

out books, or the speed with which reference questions are answered. Rather, they want libraries to determine what students know and are able to do as a result of their interaction with the library and its staff.

In the past two decades, academic librarians have embraced the concept that libraries contribute to college student learning and development. They have identified the skills students need to be successful: the abilities to locate, evaluate, and use information. They have communicated the importance of "information literacy" to their campuses, and they have created and deployed information literacy programs to teach students these skills. Now, librarians are charting unknown territory—assessing the impact of the programs they have launched.

To many librarians, the mandate for information literacy assessment is a crisis; they need to measure information literacy skills but feel unprepared to do so. Thus, librarians require a map to navigate through the dangers and opportunities inherent in assessment approaches. This article describes three major assessment approaches: fixed-choice tests, performance assessments, and rubrics. It maps the theoretical and educational assumptions on which these options are grounded and charts the dangers and opportunities of each.

> **To many librarians, the mandate for information literacy assessment is a crisis; they need to measure information literacy skills but feel unprepared to do so.**

Fixed-Choice Tests

In academic libraries, tests are used to "provide information about the students' library skills before and after a sequence of library instruction and/or research activities."[1] Examples of "fixed-choice" information literacy tests (multiple-choice, matching, true/false) abound in library instruction literature. Lawrie Merz and coauthor Beth Mark and Diana Shonrock published collections of information literacy tests.[2] Among the most well-known information literacy tests are those used at several University of California campuses, the Bay Area Community Colleges, and James Madison University. A few information literacy tests strive to be "standardized tests," or tests designed to be administered uniformly and scored objectively. One example is the SAILS (Standardized Assessment of Information Literacy Skills) test developed at Kent State University.

Theoretical Background

In general, fixed-choice tests are assessment measures derived from behaviorist theories of learning and educational measurement (see figure 1). In the early 1900s, principles of scientific measurement were applied to schools. Educators believed that learning tasks should be broken down into fundamental building blocks, which instructors would teach, students would learn, and instructors would measure.[3] To support this model, educators sought "precise standards of measurement…to ensure that each skill was mastered at the desired level."[4]

Early models and theories of learning and measurement have had long-lasting impact on educational assessment methods, including six key assumptions from behaviorist theories that affect beliefs about teaching and testing:

1. Learning occurs by accumulating atomized bits of knowledge.
2. Learning is tightly sequenced and hierarchical.
3. Transfer is limited, so each objective must be explicitly taught.
4. Tests should be used frequently to ensure mastery before proceeding to the next objective.
5. Tests are isomorphic with learning (tests = learning).
6. Motivation is external and based on positive reinforcement of many small steps.[5]

These assumptions continue to have an impact on current beliefs about student learning and assessment methods. They also influence beliefs about fairness in testing, the need for separation of testing from teaching, the importance of uniform administration, and the nature of "objectivity."[6]

Benefits ("Opportunities")

In addition to their deep roots in early educational theory, fixed-choice tests offer a number of benefits (see table 1). Indeed, librarians have "realized some success with…tests."[7] As quantitative measures, tests provide data in numerical form and are excellent choices for finding answers to questions of how much or how many.[8] They are easy to score and require less time and money, especially if computers are used for scoring.[9] In this way, they allow for the collection of a lot of data quickly.[10] Tests are good tools for measuring students' acquisition of facts and can be used to compare pre- and posttest results or to compare groups of students to each other.[11]

Another advantage of fixed-choice tests is that they can be made highly reliable.[12] In fact, high reliability is one of the most frequently cited advantages of tests.[13] One way to increase the reliability of an assessment tool is to make it longer. Lengthening a test is much easier than lengthening other types of assessment methods.[14] Test/re-test and parallel forms reliability estimates are easier to obtain with tests than other assessment methods that take more time or are difficult to repeat exactly.[15] Indirect assessments like fixed-choice tests also tend to have a higher predictive validity with a variety of measures, including college GPA and standardized test scores.[16]

Yet another advantage to using tests for assessment is that people believe in them.[17] Because the public is familiar with fixed-choice tests—especially standardized tests—and believes them to be extensively developed, tests can be used for "enhanced political leverage."[18] Administrators may support standardized tests because they compare students' achievement against other groups or national profiles.[19] Parents and students might also value such normative comparisons and try to use them to identify individual students' strengths and weaknesses.[20]

According to Joseph Prus and Reid Johnson, when locally developed, tests have several additional benefits. First, they have the benefit of being adapted to local goals and student characteristics. The process of developing test questions can help librarians determine what they really want to know about student learning. Local grading is an additional benefit—librarians have control over the interpretation and use of the results, and students receive immediate feedback. Non-locally developed standardized tests also offer two benefits. They can be implemented quickly and they can reduce the staff time required for development and scoring.[21]

Table 1
Benefits and Limitations of Fixed-Choice Tests

Considerations	Benefits "Opportunities"	Limitations "Dangers"
Learning	• Measure acquisition of facts	• Measure recognition rather than recall • Do not test higher-level thinking skills • Do not measure complex behavior or "authentic" performances • Do not facilitate learning through assessment • Include oversimplifications • Reward guessing
Data	• Are easy and inexpensive to score • Provide data in numerical form • Collect a lot of data quickly • Tend to have high predictive validity with GPA or standardized test scores • Can be made highly reliable • Can be used to compare pre- and post-tests • Can be used to compare groups of students	• May be designed to create "score-spread" • May be used as "high stakes" tests
If locally developed…	• Are adapted to local learning goals and students • Help librarians learn what they want to know about student skills	• Difficult to construct and analyze • Require leadership and expertise in measurement

Considerations	Benefits "Opportunities"	Limitations "Dangers"
	• Can be locally graded and interpretation of results can be controlled	• May not be useful for external comparisons
If non-locally developed…	• Can be implemented quickly • Reduce staff time required for development and scoring	
Other	• Are widely accepted by general public	

Limitations ("Dangers")

In the past, tests served as reasonable indicators of student learning. Now, many educators believe that traditional tests are "inadequate and misleading as measures of achievement."[22] Although fixed-choice tests are widely used and, in some arenas, still

widely respected, other educators consider them "fallible, limited measures of learning goals."[23] Indeed, tests have several possible limitations. They do not assess higher-level thinking skills, they lack authenticity, and they tend to have the problems associated with all high-stakes testing. They are time consuming to create, difficult to analyze,

> **A major limitation of fixed-choice tests is that they are indirect assessments that fail to measure higher order thinking skills.**

and problematic on a local level. Because many librarians are unfamiliar with the constraints of fixed-choice tests, a detailed description of these limitations is warranted.

A major limitation of fixed-choice tests is that they are indirect assessments that fail to measure higher order thinking skills.[24] As "objective" tests, they measure low-level recognition rather than recall.[25] Because of artificial time limits and the pressure to survey as much content as possible, fixed-choice tests rarely involve multiple dimensions of the same topic.[26] By focusing only on individual parts of a concept, test creators tend to develop over-simplified test items. In fact, the fixed-answer choices limit the ability to measure changes in "complex behavior or actual performance success."[27] Because of these limitations, it is difficult to use tests to measure the results of improved information literacy instruction.[28] Furthermore, indirect assessments like fixed-choice tests may "dramatically under-represent" constructs like information literacy.[29] Also, tests do not necessarily help students learn and develop complex skills, which should ultimately be a goal of good assessment. Esther Grassian and Joan Kaplowitz state that such tests are "less valid for testing higher-level cognitive skills such as analysis, synthesis, and

evaluation, or to determine process learning and the acquisition of concepts. As such, they may not be appropriate for many of [librarians'] needs."[30]

A second limitation of fixed-choice tests is their inability to provide an authentic assessment of student learning. Tests typically create "an artificial situation that does not really test how the learner would react in a real-world situation."[31] Therefore, they sacrifice authenticity.[32] One problem is that the conditions of fixed-choice tests are highly controlled. Students must work within time limits, with limited access to resources, and with few opportunities to make revisions.[33] Consequently, tests tend to "overassess student 'knowledge' and underassess student 'know-how with knowledge.'"[34] As a result, students who score well on fixed-choice tests may only be demonstrating that they are good test-takers. When faced with a real-world scenario, these students may not be able to formulate an appropriate response.[35] This is a dangerous limitation of fixed-choice tests because it may signal to students that the point of learning is not to "acquire 'useable knowledge,' but rather 'testable knowledge.'"[36]

A third important limitation of fixed-choice tests is that most are designed to produce variance of scores, or "score-spread," to allow comparisons among students or groups of students. According to W. James Popham, "It is this quest for score spread that renders such tests unsuitable" for measuring program and instructor effectiveness.[37] Furthermore, the time constraints of most standardized tests exacerbate the problem of score-spread. Time constraints are necessary so that students do not become "restless or, worse, openly rebellious," but they cause test designers to strive for maximum score-spread in the fewest number of test items.[38] As a result, test designers include a large number of items that only 40 to 60 percent of the students will answer correctly, as these are the questions that will produce the most score-spread. This means that very few of the questions that 80 to 90 percent of the students will answer correctly—the questions that cover the content most emphasized by instructors—remain on the test. Consequently, fixed-choice tests do not adequately cover the content that teachers consider most important, and, therefore, they are not useful in detecting effective instruction.[39]

Because the problems associated with "score-spread" are well known, some educators call for a greater emphasis on "criterion-based" tests. Criterion-based tests reference a score back to a defined skill, standard, or body of knowledge, not to other students' performances. However, truly criterion-referenced tests are elusive. Tests that purport to be criterion-referenced are usually constructed in the same way as traditional standardized tests.[40] According to Johanna Crighton:

> In practice,…when constructing criteria for a criterion-referenced test, norm-referencing is unavoidable. Hidden behind each criterion is norm-referenced data: assumptions about how the average child in that particular age group can be expected to perform. Pure criterion-referenced assessment is rare and it would be better to think of assessment as being a hybrid of norm- and criterion-referencing. The same is true of setting standards, especially if they have to be reachable by students of varying ability: one has to know something about the norm before one can set a meaningful standard.[41]

Indeed, tests are not inherently norm-referenced or criterion-referenced. What is norm-referenced or criterion-referenced is the inferences educators make from the results.[42] Furthermore, to be a meaningful assessment option, criterion-referenced tests

must be based on standards that are explicit and measurable. Oftentimes, standards are written in unmeasurable terms.[43] To serve as a basis for criterion-based tests, standards must articulate the behaviors to be measured "with sufficient clarity so that the teacher can teach toward the bodies of skills and knowledge the tests represent. …Content standards assessed must be accompanied by an assessment description that sets forth…just what a student needs to be able to do cognitively to perform well," complete with illustrative examples.[44] This level of description is rarely seen in learning standards created by professional associations.

A fourth significant limitation of fixed-choice tests is that they share characteristics of "high-stakes" tests; and, in some cases, they are actually used as high-stakes tests. By definition, high-stakes tests are tests used to make determinations affecting the future of a student, instructor, or program.[45] The "Standards for Educational and Psychological Testing" state, "Individual students should be protected from tests being used as the sole criterion for critically important decisions."[46] Many educators agree that standardized testing approaches should not be used as a basis for promotion into or exclusion from an instructional program. Tests used to make such decisions should "meet the most stringent technical standards because of the harm to individuals that would be caused by test inaccuracies."[47] This is problematic because group-administered multiple-choice tests "always include a potentially higher degree of error, usually not correctable by 'guessing correction' formulae," resulting in lower test validity.[48] For these reasons, high-stakes tests should not be used as sole determiners of major educational decisions for individual students.

High-stakes testing has additional negative implications for teaching and learning. According to Lori Shepard, high-stakes tests lead to score inflation and curriculum distortion.[49] Shepard continues, "Under intense political pressure, test scores are likely to go up without a corresponding improvement in student learning. In fact, distortions in what and how students are taught may actually decrease students' conceptual understanding."[50] She notes that teachers will "teach to the test" when student scores have serious consequences.[51]

Locally developed fixed-choice tests carry additional limitations. To begin with, the process of constructing and analyzing a test is difficult. Locally developed tests require leadership, coordination, and expertise in measurement.[52] Furthermore, a test with good psychometric properties can take years to develop.[53] As a result, locally developed tests may not provide for "externality" or a "degree of objectivity associated with review and comparisons external to the program or institution."[54] Even if a locally developed assessment is deemed adequate, it is important not to administer the same test so often that students become "over-surveyed."[55]

Performance Assessments

Because of the limitations of fixed-choice tests, the quantitative methods that once dominated assessment in higher education are slowly being replaced by qualitative forms of assessment that require students to perform real-life applications of knowledge and skills.[56] These "performance assessments" reinforce the concept that what students learn in class should be usable outside the classroom.[57] Lesley Farmer explains, "It's

the difference between describing how to ride a bike and actually putting the foot to the petal [sic] and pumping down the street. Thus a scantron 'bubble' test would be an unlikely…assessment tool."[58]

Performance assessments are based on beliefs about teaching that "differ significantly from assumptions of the past."[59] In the performance assessment paradigm, learning is an active process in which "students construct meaning and knowledge: they do not have meaning or knowledge handed to them in a book or lecture. Learning, then, is a process of students 'making sense' of how things fit together; factual and procedural knowledge is built along the way."[60] To measure learning, performance assessments give students "opportunities to demonstrate their understanding and to thoughtfully apply knowledge, skills, and habits of mind in a variety of contexts" as they would in the real world.[61]

There are a number of ways to structure performance assessment. A performance assessor might observe students as they perform a task or examine the products that result from a task performance and judge their quality.[62] For example, several authors describe assessments based on the analysis of bibliographies and paper citations.[63] Loanne Snavely and Carol Wright and Ina Fourie and Daleen Van Niekerk report their experiences with portfolio assessment.[64] Mark Horan describes the use of sketch maps in a creative approach to information literacy assessment.[65] In addition to these performance assessments, the Educational Testing Service has recently developed a computer-based performance test of information and communication technology skills called the iSkills test.

Performance assessments should meet a number of goals. According to Grant Wiggins, performance assessments should be meaningful, authentic, and involve actual "performances, not drills."[66] Performance assessments should "simulate as much as possible the situations in which students would make integrated use of [new] knowledge, skills, and values."[67] Performance assessments should require "complex and challenging mental processes from students. They should acknowledge more than one approach or one right answer and should place more emphasis on un-coached explanations and real student products."[68] Finally, they should be open-ended enough to allow each student "to bring to it his individual gifts and to maximize individual learning."[69]

Theoretical Background: Education

Performance assessment is grounded in constructivist educational theory. According to James Elmborg, "The essential defining trait of [constructivist] theorists is an insistence that knowledge is 'constructed' by individuals rather than passed on fully-formed from teachers to students."[70] Thus, constructivist theories of education are in opposition to earlier models that are firmly entrenched in educational traditions.[71] In constructivist environments, learning is promoted through active engagement and purposeful interaction in real world, authentic problem solving, critical thinking, and knowledge creation.[72] Learning, in this way, is "integrated and complex" rather than "sequential and linear."[73] Constructivist instructors help learners connect new knowledge to old knowledge, act as facilitators, guide students through a process of "cognitive restructuring" rather than memorizing facts, involve students in teaching and learning processes, and continually reflect on their practices.[74] This constructivist approach to teaching and learning is now widely accepted by most educational researchers.[75]

Theoretical Background: Motivation

Not only does performance assessment suit the requirements of constructivist pedagogy but it is also well aligned with educational motivation theories. Educators and assessors concerned with student motivation face an uphill battle. Mark Battersby strongly states that many students in higher education have an alienated attitude about learning and think it is "something you do in school, for school."[76] He suggests that students

> are often motivated not by the desire for enhanced understanding and intellectual powers, but rather to satisfy (what they often see as arbitrary) requirements. We do not believe that any teacher sets out to induce such superficial learning in students, but all too often curricular traditions, pedagogy, or school structure encourage students to adopt just such an "alienated" posture towards their learning.[77]

Student alienation can be exacerbated by assessments that use extrinsic rewards to motivate students. In fact, overly powerful extrinsic rewards can counteract students' intrinsic motivation to learn.[78] Shepard writes:

> Established findings from the motivational literature have raised serious questions about test-based incentive systems. Students who are motivated by trying to do well on tests, instead of working to understand and master the material, are consistently disadvantaged in subsequent endeavors. They become less intrinsically motivated, they learn less, and they are less willing to persist with difficult problems.[79]

Performance assessments offer one way to minimize external rewards and emphasize students' intrinsic motivation. According to Battersby, students are motivated by educational projects that "enabl[e] them to lead a richer and more empowered life rather than…a task done primarily to satisfy the demands of others (passing the test)."[80] Performance assessments fit this description. They provide "learner-centered, problem-driven approaches…[and] are…effective in engagement, motivation, and, through their problem-driven format, in providing a solid conceptual understanding."[81]

> **Performance assessments offer one way to minimize external rewards and emphasize students' intrinsic motivation.**

Theoretical Background: Assessment

In addition to educational and motivational theories, performance assessment is grounded in "assessment for learning" theories. Assessment for learning theory suggests that "good teaching is inseparable from good assessing,"[82] and assessment should be thought of not just as evaluation, but as a "primary means" of learning that requires the use of meaningful and complex assessment assignments.[83] Judith Arter agrees that assessments can be tools for learning and that students should learn by completing an assessment. She explains, "Educators do not teach and then assess; nor do they think of assessment as something that is done to students. Instead, they consider the assessment activity itself an instructional episode."[84]
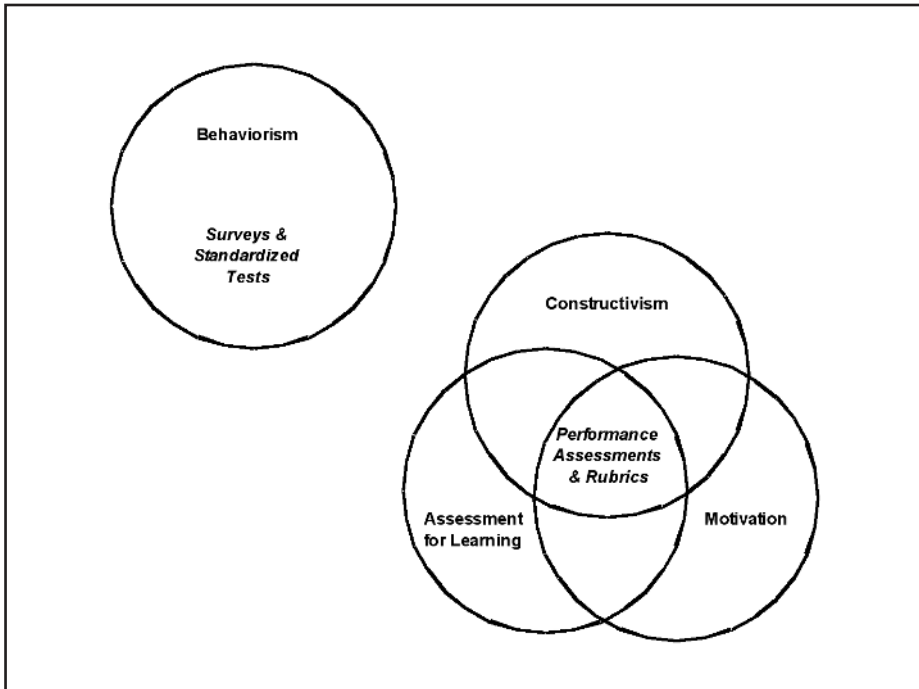
Figure 1. Theory Underlying Assessment Approaches

Ideally, the assessment task should be one that is already inherent in a curriculum or learning program. If an embedded task cannot be found, then the assessment task should be one that is easily included in day-to-day instruction.[85] Shepard recommends assessments that provide opportunities for students to demonstrate their ability to connect concepts, recognize problem solving patterns, and monitor their own learning.[86] Assessment tasks that offer all of these qualities are easily found among regular classroom learning tasks.[87]

Benefits ("Opportunities")

Performance assessment offers numerous benefits to teachers and learners (see table 2). Among them are (1) close connections between instruction and assessment, (2) ability to measure higher-order thinking skills, (3) contextualization of assessment that leads to transfer of knowledge, greater equitability, and increased validity, and (4) ability to use results to improve instruction and programs.

Perhaps the greatest value of performance assessment is that the form and content of the assessment method can be closely aligned with instructional goals.[88] As a result of this alignment, "the use of performance assessment in the classroom has been seen by some as a promising means of accomplishing a long-standing, elusive goal—namely, the integration of instruction and assessment."[89]

Performance assessments allow educators to capture the higher-order thinking and reasoning skills typically absent in more traditional forms of assessment.[90] Lauren

Resnik describes the ways in which performance tasks focus on higher-order skills. She states that higher-order thinking is complex in that the "total path of action is not 'visible' (mentally speaking) from any single vantage point," yields multiple solutions, and requires nuanced judgment.[91] Higher-order skills require students to apply multiple criteria, deal with uncertainty and the unknown, regulate their own thinking, and impose meaning on structures that initially seem meaningless.[92] All of these higher-order thinking processes are present in performance assessments.

Another benefit of performance assessments is that they are contextualized. According to Grant Wiggins, "There is no such thing as performance-in-general. To understand what kind and precision of answer fits the problem at hand, the student needs contextual detail."[93] Performance assessments recognize the contexts in which students work and aim to "invent an authentic simulation,…and like all simulations…the task must be rich in contextual detail."[94] Through contextualization, performance assessments help students understand the relevance of what they learn.[95] Performance assessments also can "reflect…society's demands for education that prepares students for the world of work."[96]

The contextualization of performance assessment offers three more benefits: transfer of knowledge, equitability, and validity. By contextualizing assessment in real world problems, performance assessments allow students to transfer knowledge and use it in new ways.[97] Using contextualized, authentic assessment is a more equitable approach, as opposed to using tests susceptible to the "bias associated with testing rapid recall of decontextualized information."[98] Finally, performance assessments may be able to render more valid data than other types of assessments. David Sweet points out that "there is a big difference between answering multiple-choice questions on how to make an oral presentation and actually making an oral presentation."[99]

Limitations ("Dangers")

A few technical and feasibility issues have, in the past, thwarted attempts to use performance assessment on a large scale. Among these are cost, time, and generalizability. Compared to traditional tests, performance assessments can be costly to develop, administer, and score.[100] For example, the ETS iSkills performance assessment is the most expensive information literacy assessment available for purchase. Time is a second limitation of performance assessments. They require greater time, in both planning and thought, from both teachers and students.[101] Teachers must select observable outcomes, design authentic performances to capture evidence of student learning, collect artifacts from student performances, and analyze them to evaluate student learning. Students must demonstrate behaviors or provide products that show their behaviors; both options generally require more time than completing fixed-choice tests. A third limitation of performance assessments is that results may not be generalizable to other settings and populations.[102] To achieve generalizable results, multiple assessments may be required.[103] This is a significant concern for libraries that want to benchmark their assessments against peer institutions. Limitations of cost, time, and generalizability should not be underestimated; however, in the future, advances in cognitive sciences and assessment technology may resolve these three limitations of performance assessment.[104]

Table 2
Benefits and Limitations of Performance Assessments

Considerations	Benefits "Opportunities"	Limitations "Dangers"
Learning	• Capture higher-order thinking skills • Align with learning goals • Integrate learning and assessment • Support learning and assessment in authentic contexts • Facilitate transfer of knowledge	
Data	• Supply valid data	• May have limited generalizability to other settings and populations
Other	• Offer equitable approach to assessment	• Are costly to create, administer, and score

Rubrics

Many librarians who assess information literacy instruction are familiar with fixed-choice tests; some have investigated the use of performance assessments. Comparatively few have explored the use of rubrics, despite the fact that assessment experts proclaim rubrics "one of the most basic tools" for measuring learning.[105] In fact, only a handful of authors report the use of rubrics to assess information literacy skills.[106]

> **Comparatively few have explored the use of rubrics, despite the fact that assessment experts proclaim rubrics "one of the most basic tools" for measuring learning.**

Rooted in constructivist, motivation, and assessment for learning theories, rubrics are "descriptive scoring schemes" created by educators to guide analysis of student work.[107] John Haffner writes, "In the educational literature and among the teaching and learning practitioners, the word 'rubric' is understood generally to con-

note a simple assessment tool that describes levels of performance on a particular task and is used to assess outcomes in a variety of performance-based contexts."[108] Wiggins defines rubrics as

> a set of scoring guidelines for evaluating students' work. Rubrics answer the following questions: By what criteria should performance be judged? Where should we look and what should we look for to judge performance success? What does the range in the quality of performance look like? How do we determine validly, reliably, and fairly what score should be given and what that score means? How should the different levels of quality be described and distinguished from one another?[109]

Rubrics are often employed when educators must judge the quality of performances or constructed-response items,[110] and they can be used across a broad range of subjects.[111]

Benefits ("Opportunities")

The benefits of rubric assessment are numerous for both students and librarians (see table 3). For students, rubric advantages include: (1) students understand the expectations of their instructors, (2) students are empowered to meet standards, (3) students receive direct feedback about current and future learning, and (4) student self-evaluation is facilitated. For librarians engaged in instructional assessment, rubric advantages include: (1) agreed upon values, (2) reliable scoring of student work, (3) detailed result data, (4) a focus on standards-based education, (5) evaluation of student learning across time or multiple programs, and (6) cost.

Rubrics benefit students in several ways. First, rubrics help students learn more effectively[112] because they state instructional expectations clearly,[113] and "people usually gain more from a lesson when they are told what it is they are supposed to learn."[114] Lois Pausch and Mary Popp suggest that rubrics are more valuable for learners than other assessment tools because they emphasize "understanding rather than memorization, 'deep' learning rather than 'surface' learning."[115]

Because rubrics provide criteria and quality standards, they can be used to empower students. John Herman, Pamela Aschbacher, and Lynn Winters describe the importance of discussing scoring rubrics with students:

> Public discussions help students to internalize the standards and "rules" they need to become independent learners. …Examples of what constitutes good work engage students in the work itself and in judgments about their work. Public discussions of quality and criteria inform students during the formative period of instruction, not simply at the end of a unit or course when it is too late to make improvements. Furthermore, discussions of criteria also help students see the perspectives of their teachers, their peers, and sometimes even of the experts in the field.[116]

Many educators believe that students should know the rules for how their products and performances will be judged. According to Shepard, providing students with access to evaluation criteria is required for "basic fairness."[117]

Rubrics offer a third benefit; they provide timely[118] and detailed feedback to students.[119] According to Marilee Bresciani, Carrie Zelna, and James Anderson rubrics make

rankings, ratings, and grades more meaningful by revealing what educators expect students to know or do. When rubrics are used as tools for feedback, "students begin to understand what it is they are or are not learning, and are or are not able to demonstrate what they know and can do. When students begin to see what they are not learning, they can take more responsibility for their learning."[120] Over time, rubric feedback can "provide a benchmark from one lesson to the next."[121]

As students become comfortable with rubric assessment, they can use rubrics independently for peer- or self-evaluation.[122] This process helps "students identify their own learning."[123] When students participate in self-assessment, they think critically about learning,[124] benefit from increased responsibility for learning, form a more collaborative relationship with their teachers, gain ownership, become honest about their work, grow to be more interested in learning than grades, and can be held to higher standards, all of which result in increased investment in learning.[125]

In addition to student benefits, rubrics provide a number of advantages for librarians conducting instructional assessment. First, the rubric creation process gives librarians the opportunity to discuss and determine agreed upon values of student learning. Values that librarians, classroom instructors, and students share can be communicated to library administrators, external stakeholders, and the public.[126] In fact, once rubrics are developed, they can be used to normalize stakeholders' expectations and to bring them in line with librarians' vision for student learning.[127]

Second, using rubrics that are based on agreed upon values has practical advantages for consistent scoring. When used correctly, "rubrics come close to assuring that inadequate, satisfactory, and excellent mean the same thing on the same skill set from one group of students to a similar group regardless of who makes the evaluation."[128] Rubrics that can be used consistently by multiple assessors allow instructors from different units to work together to improve instructional programs.[129]

Rubric assessment offers a third instructional benefit—data full of rich description. Rubrics provide "evaluators and those whose work is being evaluated with rich and detailed descriptions of what is being learned and what is not."[130] This level of detail also prevents inaccuracy of scoring and bias.[131] Because rubrics are easy to use and easy to explain, they generate data that are easy to understand, defend, and convey.[132] Furthermore, the data that results from rubric assessment are so detailed and well defined that "accusations that evaluators do not know what they are looking for" are combated.[133] Ultimately, this descriptive data can be used to document how librarians or administrators can improve instruction.[134]

Additional benefits of a rubric approach to assessment include a focus on standards, transferability across time or multiple programs, and cost. According to Bresciani, Zelna and Anderson, rubrics teach students "the standards of the discipline or the standards of the…learning…experience,"[135] and they help determine the extent to which students achieve standards.[136] Rubrics are flexible and can be used to assess students, activities, or programs. Educators can use the same or similar rubrics over the course of time, even an academic career, to assess student progress toward learning goals. Rubrics can also be used to assess individual students, all students in a program, or students across multiple programs.[137] This is important for educators and program administrators who need to capture the learning of a student population that flows from program to program.[138] A

final benefit of rubric assessment is cost. Because most rubrics are currently designed and implemented in-house, the main cost associated with rubrics is an investment in time for conversation and norming, rather than dollars.

Limitations ("Dangers")

Like other assessment approaches, there are limitations associated with rubric assessment. Many of the limitations of a rubric approach to assessment are rooted in poor rubric construction. Not all rubrics are well written, and crafting a good rubric requires time, practice, and revision.[139] Robin Tierney and Marielle Simon caution that, unfortunately, "the most accessible rubrics, particularly those available on the Internet, contain design flaws that not only affect their instructional usefulness, but also the validity of their results."[140] Librarians who seek to develop usable rubrics should seek a balance between generalized wording, which increases the amount of different tasks to which the rubric can be applied, and detailed description, which provides a level of specificity that ensures greater reliability; should be concrete, yet concise avoid jargon; should use parallel language; differentiate performance levels; should emphasize quality, rather than quantity; and avoid describing low levels of student performance in negative terms.[141]

Another limitation of rubric assessment is time. Although creating rubrics is relatively inexpensive monetarily, some educators consider the process time-consuming.[142] Part of that perception might be due to lack of familiarity or expertise.[143] Initially, librarians may require more training time to use rubrics effectively than classroom instructors.[144] For example, librarians must be calibrated, or "normed," on a rubric. The norming process requires multiple steps and is best completed in a face-to-face training session. Typically, norming a rubric requires a few hours of training in addition to the time spent scoring student learning samples. In addition, librarians may have several barriers that interfere with effective rubric use, and this should be addressed by training.[145] As a result, librarians seeking to use a rubric approach to assessment should invest adequate time to ensure quality results. Even so, as Joseph Prus and Reid Johnson write: "As in virtually all other domains of human assessment, there is a consistently inverse correlation between the quality of measurement methods and their expediency; the best methods usually take longer and cost more."[146]

Conclusion

Librarians seeking to assess information literacy skills have three main options. First, they may choose to purchase or design a fixed-choice test comprised of multiple-choice, matching, or true/false questions. Second, librarians may use a performance assessment approach. Librarians who choose this option may use ETS iSkills test, or they may develop their own performance assessments based on observation of student behavior or examination of student work samples. Third, librarians may choose to adopt a rubric approach to assessment. For this option, librarians must create their own rubric or adapt a preexisting one, and then apply the rubric to artifacts of student learning.

> **Certainly, no two academic libraries are the same; likewise, no two libraries have identical assessment needs.**

Table 3
Benefits and Limitations of Rubrics

Considerations	Benefits "Opportunities"	Limitations "Dangers"
Learning	• Articulate and communicate agreed upon learning goals • Focus on deep learning and higher-order thinking skills • Provide direct feedback to students • Facilitate peer- and self-evaluation • Make scores and grades meaningful • Can focus on standards	
Data	• Facilitate consistent scoring • Prevent inaccuracy of scoring and bias • Deliver data that is easy to understand, defend, and convey • Offer detailed descriptions necessary for informed decision-making • Can be used over time or across multiple programs	
Other	• Are inexpensive to design and implement	• May contain design flaws that impact data quality • Require time for development • Require time for training multiple raters

Certainly, no two academic libraries are the same; likewise, no two libraries have identical assessment needs. As a result, each library must plot a separate course to successful information literacy assessment. To do so, librarians must have a firm grasp of the theories that underlie current assessment options, the dangers and opportunities provided by each assessment approach, and the needs of their institutions and libraries. By forming mental maps of the benefits and limitations of information literacy assessment options, librarians will be able to provide evidence that their instructional efforts are effective—that students are in fact able to locate, evaluate, and use information. Only then will librarians be prepared to move confidently into the future, teaching and assessing information literacy instruction both purposefully and competently.

Megan Oakleaf is assistant professor, School of Information Studies, Syracuse University, Syracuse, NY; she may be contacted via e-mail at: moakleaf@syr.edu.

Notes

1. Lorrie A. Knight, "The Role of Assessment in Library User Education," *Reference Services Review* 30, 1 (2002): 16.
2. Lawrie H. Merz and Beth L. Mark, *Clip Note #32: Assessment in College Library Instruction Programs* (Chicago: Association of College and Research Libraries, 2002); Diana D. Shonrock, *Evaluating Library Instruction: Sample Questions, Forms, and Strategies for Practical Use* (Chicago: American Library Association, 1996).
3. Lorrie A. Shepard, "The Role of Assessment in a Learning Culture," *Educational Researcher* 29, 7 (2000): 4.
4. Ibid.
5. Ibid., 5.
6. Ibid., 5–6.
7. Knight, 16.
8. Ester S. Grassian and Joan R. Kaplowitz, *Information Literacy Instruction* (New York: Neal-Schuman, 2001), 276.
9. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, *The NPEC Sourcebook on Assessment, Volume 1: Definitions and Assessment Methods for Critical Thinking, Problem Solving, and Writing* (Harrisonburg, VA: National Center for Education Statistics, 2000), 4.
10. Grassian and Kaplowitz, 281.
11. Ibid.; Knight; and Grassian and Kaplowitz, 281.
12. Dean A. Colton et al., *Reliability Issues with Performance Assessments* (Iowa City: ACT, 1997), 3.
13. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development.
14. Dean A. Colton et al., 3.
15. Ibid. 4.
16. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development.
17. Shepard, "Why We Need Better Assessments," in *A Handbook for Student Performance Assessment in an Era of Restructuring*, ed. Robert E. Blum and Judith A. Arter (Alexandria,VA: Association for Supervision and Curriculum Development, 1996), I-2: 3.
18. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development.
19. Peggy Maki, "From Standardized Tests to Alternative Methods: Assessing Learning in Education," *Change* 33, 2 (2001): 28–32.

20. W. James Popham, *Test Better, Teach Better: The Instructional Role of Assessment* (Alexandria, VA: Association for Supervision and Curriculum Development, 2003), 126.
21. Joseph Prus and Reid Johnson, "A Critical Review of Student Assessment Options," *New Directions for Community Colleges* 1994, 88 (Winter 1994): 71–2.
22. Shepard, I-2: 4.
23. Ibid., I-2: 6.
24. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 4.
25. Grassian and Kaplowitz, 281.
26. Shepard, I-2: 3.
27. Grassian and Kaplowitz, 281.
28. Molly R. Flaspohler, "Information Literacy Program Assessment: One Small College Takes the Big Plunge," *Reference Services Review* 31, 2 (2003): 139.
29. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 4.
30. Grassian and Kaplowitz, 281.
31. Ibid.
32. Robert J. Marzano, Debra Pickering, and Jay McTighe, *Assessing Student Outcomes: Performance Assessment Using the Dimensions of Learning Model* (Alexandria, VA: Association for Supervision and Curriculum Development, 1993), 11.
33. Ibid.
34. Grant Wiggins, "Creating Tests Worth Taking," in *A Handbook for Student Performance in an Era of Restructuring*, ed. Robert E. Blum and Judith A. Arter (Alexandria, VA: Association for Supervision and Curriculum Development, 1996), V–6: 3.
35. Grassian and Kaplowitz, 277.
36. Mark Battersby, *So, What's a Learning Outcome Anyway?* (1999), ED 430 661.
37. Popham, 126.
38. Ibid., 127.
39. Ibid., 127–8.
40. Ibid., 140–1.
41. Johanna V. Crighton, "Standardized Tests and Educational Policy," *Encyclopedia of Education*, ed. James W. Guthrie (New York: Macmillan Reference USA, 2003), 7: 2530.
42. Popham, 141.
43. Crighton, 2531–2.
44. Popham, 143, 145.
45. Shepard, "Standardized Tests and High-Stakes Assessment," in *Encyclopedia of Education*, ed. James W. Guthrie (New York: Macmillan Reference USA, 2003), vol.7.
46. Ibid., 7: 2536.
47. Ibid., 7: 2535.
48. Prus and Johnson, 71–2.
49. Shepard, "Standardized Tests and High-Stakes Assessment," 2536.
50. Shepard, "The Role of Assessment in a Learning Culture," 9.
51. Shepard, "Why We Need Better Assessments," I-2: 3.
52. Prus and Johnson, 72–3.
53. Marilee J. Bresciani, Carrie L. Zelna, and James A. Anderson, *Assessing Student Learning and Development: A Handbook for Practitioners* (National Association of Student Personnel Administrators, 2004), 70.
54. Prus and Johnson, 72–3.
55. Bresciani, Zelna, and Anderson, 71.
56. Lois M. Pausch and Mary Pagliero Popp, "Assessment of Information Literacy: Lessons from the Higher Education Assessment Movement" (paper presented, ACRL 8th National Conference, Nashville, Tennessee, April 1997), http://ala.org/ala/acrlbucket/nashville1997pap/pauschpopp.cfm (accessed April 29, 2008).

57. Battersby.

58. Lesley S. Johnson Farmer, "Authentic Assessment of Information Literacy through Electronic Products," *Book Report* 16 (September/October 1997).

59. Richard J. Shavelson and Gail P. Baxter, "Linking Assessment with Instruction," in *A Handbook for Student Performance in an Era of Restructuring*, ed. Robert E. Blum and Judith A. Arter (Alexandria, VA: Association for Supervision and Curriculum Development, 1996), IV–7: 1.

60. Ibid..

61. Marzano, Pickering, and McTighe, 13.

62. Edward A. Silver, "Performance Assessment," in *Encyclopedia of Education*, ed. James W. Guthrie (New York: Macmillan Reference USA, 2003), 1: 134.

63. Virginia E. Young and Linda G. Ackerson, "Evaluation of Student Research Paper Bibliographies: Refining Evaluation Criteria," *Research Strategies* 13, 2 (1995): 80–93; Karen Hovde, "Check the Citation: Library Instruction and Student Paper Bibliographies," *Research Strategies* 17 (Spring 2000): 3–9; Flaspohler; and Lara Ursin, Elizabeth Blakesley Lindsay, and Corey M. Johnson, "Assessing Library Instruction in the Freshman Seminar: A Citation Analysis Study," *Reference Services Review* 32, 3 (2004): 284–92.

64. Loanne L. Snavely and Carol A. Wright, "Research Portfolio Use in Undergraduate Honors Education: Assessment Tool and Model for Future Work," *Journal of Academic Librarianship* 29, 5 (2003): 298–303; Ina Fourie and Daleen Van Niekerk, "Using Portfolio Assessment in a Module in Research Information Skills," *Education for Information* 17, 4 (1999): 333–53.

65. Mark Horan, "What Students See: Sketch Maps as Tools for Assessing Knowledge of Libraries," *Journal of Academic Librarianship* 25, 3 (1999): 187–201.

66. Wiggins, V–6: 3, 4.

67. David Sweet, "Performance Assessment," *Education Research Consumer Guide* 2 (Office of Research, Office of Educational Research and Improvement (OERI), U.S. Department of Education, 1993), http://www.ed.gov/pubs/OR/ConsumerGuides/perfasse.html (accessed April 29, 2008).

68. Shepard, "Why We Need Better Assessments," I-2: 4.

69. Farmer.

70. James K. Elmborg, "Teaching at the Desk: Toward a Reference Pedagogy," *portal: Libraries and the Academy* 2, 3 (2002): 457.

71. Mary Lamon, "Constructivist Approach," in *Encyclopedia of Education*, ed. James W. Guthrie (New York: Macmillan Reference USA, 2003), 4: 1463.

72. Ina Claire Gabler and Michael Schroeder, *Seven Constructivist Methods for the Secondary Classroom: A Planning Guide for Invisible Teaching* (Boston: Allyn and Bacon, 2003), 4; Chandra J. Foote, Paul J. Vermette, and Catherine F. Battaglia, *Constructivist Strategies: Meeting Standards and Engaging Adolescent Minds* (Larchmont, NY: Eye on Education, 2001), 24.

73. Ibid.

74. Gabler and Schroeder, 4–5.

75. Lamon, 1466.

76. Battersby.

77. Ibid.

78. Mark R. Lepper and Jennifer Henderlong, "Motivation: Instruction," in *Encyclopedia of Education*, ed. James W. Guthrie (New York: Macmillan Reference USA, 2003), 5:1696.

79. Shepard, "Standardized Tests and High-Stakes Assessment," 2536.

80. Battersby,

81. Donald A. Norman and James C. Spohrer, "Learner-Centered Education," *Communications of the ACM* 39, 4 (1996): 26–7.

82. Wiggins, V–6: 8.

83. Battersby.

84. Judith A. Arter, "Using Assessment as a Tool for Learning," in *A Handbook for Student Performance in an Era of Restructuring*, ed. Robert E. Blum and Judith A. Arter (Alexandria, VA: Association for Supervision and Curriculum Development, 1996), IV–10: 1.

85. Bresciani, Zelna, and Anderson, 22.

86. Shepard, "Why We Need Better Assessments," I-2: 4.

87. Ibid.

88. Silver, 135.

89. Ibid.

90. Ibid.

91. Lauren B. Resnick, *Education and Learning to Think* (Washington, D.C.: National Academy Press, 1987), 3.

92. Ibid.

93. Wiggins, V–6: 5.

94. Ibid., V–6: 3.

95. Farmer.

96. Ibid.

97. Shepard, "The Role of Assessment in a Learning Culture," 11.

98. Silver, 135.

99. Sweet.

100. Silver, 135.

101. Sweet.

102. Silver, 135.

103. Popham, 103.

104. Silver, 136.

105. Grant Wiggins, *Educative Assessment: Designing Assessments to Inform and Improve Student Performance* (San Francisco: Jossey-Bass, 1998), 153.

106. Barbara J. D'Angelo, "Integrating and Assessing Information Competencies in a Gateway Course," *Reference Services Review* 29, 2 (2001): 282–93; Merz and Mark; Ilene Rockman, "Rubrics for Assessing Information Competence in the California State University" (October 2002), http://www.calstate.edu/LS/1_rubric.doc (accessed April 16, 2008); Mark Emmons and Wanda Martin, "Engaging Conversation: Evaluating the Contribution of Library Instruction to the Quality of Student Research," *College and Research Libraries* 63, 6 (2002): 545–60; Elizabeth Fuseler Avery, *Assessing Student Learning Outcomes for Information Literacy Instruction in Academic Institutions* (Chicago: Association of College and Research Libraries, 2003); "Bay Area Community Colleges Information Competency Assessment Project" (2003), http://www.topsy.org/ICAP/ICAProject.html (accessed April 16, 2008); Lorrie A. Knight, "Using Rubrics to Assess Information Literacy," *Reference Services Review* 34, 1 (2006): 43–55; Megan J. Oakleaf, "Using Rubrics to Collect Evidence for Decision-Making: What Do Librarians Need to Learn?" *Evidence Based Library and Information Practice* 2, 2 (2007): 27–42; and Elizabeth Choinski, Amy E. Mark, and Missy Murphey. "Assessment with Rubrics: An Efficient and Objective Means of Assessing Student Outcomes in an Information Resources Class," *portal: Libraries and the Academy* 3, 4. (2003): 563–75.

107. Barbara M. Moskal, "Scoring Rubrics: What, When, and How?" *Practical Assessment Research and Evaluation* 7, 3 (2000), http://pareonline.net/getvn.asp?v=7&n=10 (accessed April 29, 2008).

108. John C. Hafner, "Quantitative Analysis of the Rubric as an Assessment Tool: An Empirical Study of Student Peer-Group Rating," *International Journal of Science Education* 25, 12 (2003): 1509.

109. Wiggins, 154.

110. Popham, 95.

111. Moskal.

112. Popham, 95.

113. Cyndi Phillip, "Clear Expectations: Rubrics and Scoring Guides," *Knowledge Quest* 31, 2 (2002): 26.

114. Art A. Lichtenstein, "Informed Instruction: Learning Theory and Information Literacy," *Journal of Educational Media and Library Sciences* 38, 1 (2000): 28.

115. Pausch and Popp.

116. John Herman, Pamela R. Aschbacher, and Lynn Winters, "Setting Criteria," in *A Handbook for Student Performance in an Era of Restructuring*, ed. Robert E. Blum and Judith A. Arter (Alexandria, VA: Association for Supervision and Curriculum Development, 1996), VI–4: 3.

117. Shepard, "The Role of Assessment in a Learning Culture," 11.

118. Dannelle D. Stevens and Antonia Levi, *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning* (Sterling, VA: Stylus, 2005), 17.

119. Moskal.

120. Bresciani, Zelna, and Anderson, 31.

121. Daniel Callison, "Rubrics," *School Library Media Activities Monthly* 17, 2 (2000): 35.

122. Rosemarie Bernier, "Making Yourself Indispensible by Helping Teachers Create Rubrics," *CSLA Journal* 27, 2 (2004): 25.

123. Bresciani, Zelna, and Anderson, 31.

124. Stevens and Levi, 21.

125. Shepard, "The Role of Assessment in a Learning Culture," 12.

126. Bernier, 25.

127. Bresciani, Zelna, and Anderson, 31.

128. Callison, 35.

129. Bresciani, Zelna, and Anderson, 31.

130. Ibid., 30.

131. Popham, 95; Bresciani, Zelna, and Anderson, 31.

132. Heidi Goodrich Andrade, "Using Rubrics to Promote Thinking and Learning," *Educational Leadership* 57, 5 (2000): 14.

133. Bresciani, Zelna, and Anderson, 30.

134. Bernier, 25.

135. Bresciani, Zelna, and Anderson, 31.

136. Moskal.

137. Bresciani, Zelna, and Anderson, 31.

138. Ibid., 35.

139. Popham, 95; Callison, 35.

140. Robin Tierney and Marielle Simon, "What is Still Wrong with Rubrics: Focusing on the Consistency of Performance Criteria Across Scale Levels," *Practical Assessment, Research, and Evaluation* 9, 2 (2004), http://pareonline.net/getvn.asp?v=9&n=2 (accessed April 22, 2008).

141. Ibid.; Popham, 99, 73; Tierney and Simon; Callison, 36; and Tierney and Simon.

142. Ibid.

143. Bernier, 25.

144. Megan Oakleaf, "Using Rubrics to Collect Data for Decision-Making: What Do Librarians Need to Learn?" (paper, 4th International Evidence-Based Library and Information Practice Conference, Chapel Hill, NC, May 2007).

145. Oakleaf, "Using Rubrics," *Evidence Based Library and Information Practice* 2, 3 (2007): 38.

146. Prus and Johnson, 25.