

PROCEEDINGS OF THE
LIBRARY ASSESSMENT CONFERENCE
BUILDING EFFECTIVE, SUSTAINABLE, PRACTICAL ASSESSMENT

SEPTEMBER 25–27, 2006

CHARLOTTESVILLE, VIRGINIA

EDITED BY

Francine DeFranco, University of Connecticut

Steve Hiller, University of Washington

Lisa Janicke Hinchliffe, University of Illinois at Urbana-Champaign

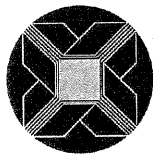
Kristina Justh, Association of Research Libraries

Martha Kyrillidou, Association of Research Libraries

Jim Self, University of Virginia

Joan Stein, Carnegie Mellon University

UNIVERSITY
of VIRGINIA
LIBRARY



ASSOCIATION OF
RESEARCH LIBRARIES

UNIVERSITY OF
WASHINGTON
Libraries

The Right Assessment Tool for the Job: Seeking a Match between Method and Need

Megan J. Oakleaf
Syracuse University, USA

Abstract

Until recently, the demands faced by other academic units on campus have passed over college and university libraries. Now, academic librarians are increasingly pressured to prove that the resources and services they provide result in improvement in student learning and development. In answer to calls for accountability, academic librarians must demonstrate their contribution to the teaching and learning outcomes of their institutions. One way librarians can achieve this goal is by assessing information literacy instruction on campus.

When assessing information literacy skills, academic librarians have a variety of outcomes-based assessment tools from which to choose. The selection of an assessment tool should be based on the purposes of an assessment situation and the "fit" between the needs of an assessment situation and the strengths and weaknesses of individual assessment approaches. In this paper, participants will learn about the major purposes of assessment and how they impact the criteria used to select an assessment tool. Among the purposes of assessment that will be discussed are the needs to respond to calls for accountability, to participate in accreditation processes, to make program improvements, and to enrich the student learning experience. Participants will also learn about several criteria useful in selecting an assessment tool including utility, relevance, stakeholder needs, measurability, and cost.

The paper will also describe the benefits and limitations of several outcomes-based approaches to the assessment of student learning outcomes. Specifically, participants will learn about assessment tools that evaluate student learning outcomes taught via information literacy instruction. The theoretical support, benefits, limitations, and relevant research related to the use of surveys, tests, performance assessments, and rubrics will be outlined and examples given of each assessment tool.

Armed with background knowledge and lessons learned from colleagues throughout higher education, academic librarians can embrace the effective and efficient assessment of information literacy instruction efforts on their campuses. This presentation will give participants a "jump start" toward becoming proficient assessors ready to answer calls for accountability.

Introduction

In the 21st century, institutions of higher education face calls for accountability. In this environment, academic librarians have responded to requests for the evaluation of information literacy skills and library instruction programs by learning about and testing assessment approaches. Before librarians launch into detailed assessment efforts, however, they should carefully consider the purpose of such assessments, use established criteria to select assessment tools, and understand the benefits and limitations of specific assessment approaches.

Purposes of Assessment

Before beginning assessment projects, librarians should first consider the purpose of their assessment efforts. Shepard states:

The intended use of an assessment—its purpose—determines every other aspect of how the assessment is conducted. Purpose determines the content of the assessment (What should be measured?); methods of data collection (Should the procedures be standardized? Should data come from all students or from a sample of students?); technical requirements of the assessment (What level of reliability and validity must be established?); and finally, the stakes or consequences of the assessment, which in turn determine the kinds of safeguards necessary to protect against potential harm from fallible assessment-based decisions.¹

Librarians who assess information literacy skills

and programs may have multiple purposes in mind. Among them are two overarching goals: 1) to respond to accountability measures, including accreditation mandates, and 2) to improve library instruction programs, and 3) increase student learning.

Responding to Calls for Accountability

Perhaps the most common purpose for assessment is to respond to a call for accountability from institutional administrators or accreditors.²

Accreditors en masse acknowledge the importance of information literacy skills. According to Gratch Lindauer, in recent years most accreditation standards have increased their emphasis on the teaching role of libraries.³ As a result, librarians increasingly find themselves at the front and center of accreditation processes.

Some authors believe that reflection initiated by accreditation may result in a significant paradigm shift in librarianship. Accreditation brings librarians the opportunity to reflect and “determine whether the library is asking the right questions, collecting useful data, analyzing the data effectively, disseminating the data to those who can benefit, and relying upon data effectively for decision making and improvement.”⁴ Thompson suggests that accreditation requirements demand:

A shift in the established library instruction paradigm at many institutions. Responsibility shifts from librarians teaching students how to locate materials for particular assignments, to faculty and librarians working together to embed the teaching and learning of information literacy skills systematically into syllabi and curricula. The new paradigm requires librarians and faculty to adopt a broader sense of the role of information literacy skills in higher education and in the preparation for the professional workforce. It also demands the learning of new methods and concepts by both teaching faculty and librarians, as they develop a collaborative approach to the integration of information literacy into general education and disciplinary education.⁵

Although accreditation processes hold the promise of rethinking and reflection among librarians, they may also cause a number of difficulties.

In many universities, the “teaching-learning role of academic libraries is well established, as are the expectations of accreditation agencies that libraries connect their evaluation of collections, resources,

and services to educational outcomes.”⁶ On other campuses, these library roles and expectations are new. As librarians reposition themselves at the center of teaching and learning (and thus accreditation) they may turn to existing measures of library effectiveness for help. Unfortunately, much of the data recorded in annual reports and reviews will not help libraries demonstrate how libraries impact student learning.⁷ Furthermore, accreditation documents do not generally outline the ways in which libraries might demonstrate the effects of their programs on institutional effectiveness.⁸ As a result, at some universities, “librarians ponder their mandate but concede a lack of authority or resources to succeed with anything ambitious.”⁹

Improving Program Performance & Structure

In addition to responding to external calls for accountability, another significant purpose for assessment is to inform decision-making¹⁰ in order to improve the quality or performance of a program.¹¹ Several authors confirm the role of assessment in program improvement. Samson writes, “Assessment offers a value-added dimension to a library instruction program. It provides a beginning point to ascertain the program’s effectiveness and to guide direction for future instruction.”¹² Without assessment, program weaknesses cannot be identified and corrected through effective decision-making. Knight acknowledges the importance of documenting both the strengths and “areas of improvement” in library instruction programs.¹³ Perhaps Barclay states the connection between assessment and improvement most clearly: “Unless evaluation will somehow improve the thing being evaluated, it is not worth doing.”¹⁴

Assessment can also have the very practical purpose of improving the structure of a program. For example, assessment can be used to connect a program mission with the missions of the larger organizational body or institution. Assessment can also be used to reinforce, emphasize, align, or realign activities with a previously defined mission of a program.¹⁵ Additionally, assessment leads to opportunities for celebrating successful achievement of a program’s mission.¹⁶ Success can be leveraged to form collaborations, find funding, and champion change. Knight asserts that the data resulting from assessments help strengthen librarian connections with faculty members.¹⁷ Assessment can also support requests to continue

or increase funding. This purpose is important for information literacy instruction programs, especially those that must justify their existence or risk losing financial support.¹⁸ Even if assessment yields negative results, it can still be used to improve program structures. Grassian and Kaplowitz suggest that librarians can use negative results to argue for greater financial support by highlighting where funding could be used to improve the program.¹⁹ As Knight suggests, "It is . . . important to view assessment programs not as ends . . . but . . . as significant sources of information that foster feedback for change."²⁰

Improving Student Learning and Teacher Skills

Although assessment is often conducted to respond to calls for accountability, participate in accreditation, or improve programs, many authors argue that the main purpose of assessment is to improve student learning. According to Popham, "The central mission of all . . . assessment is (1) to help you make valid inferences about your students so that you can then (2) make better decisions about how to instruct those students."²¹

Assessment allows programs to demonstrate their contribution to student learning and a quality university experience. The ACRL Instruction Section recognizes that library instruction programs must become involved in assessment to document the campuswide impact of library instruction and information literacy.²² Instruction librarians must "find out whether or not what is taught is useful."²³ They must demonstrate that students acquire information literacy skills through library instruction.

Assessment not only documents student learning, but also provides important feedback teachers can use to improve their skills. Grassian and Kaplowitz state that librarians, "assess, evaluate, and revise because we want to find out if our instruction has been effective. In other words, we need to find out how well our goals and objectives have been met. Furthermore, we want to highlight areas where our efforts might be improved for the future."²⁴ Sampson writes, "Assessment provides the opportunity to take a fresh look at the classroom experience."²⁵ Arter confirms that assessment "demystif[ies] the learning process" and makes teachers and learners equal stakeholders in learning.²⁶ Through assessment, teachers and students become partners in the learning process.

Criteria for Selection

Once a decision is made to conduct an assessment, the next step is to select a method or tool for assessment. Maki writes:

As institutions increasingly commit to assessing student learning . . . the first challenge they face is to decide which assessment methods to use. Identifying or developing assessment methods requires a thorough understanding of what each actually measures and how each relates to a program's articulated outcomes and expected levels of student performance. That is, what methods best enable an institution to determine how well students have met its expectations?²⁷

Rather than selecting the method that is most familiar, librarians should select assessment methods intentionally.²⁸ Among the criteria librarians should consider are: utility, relevance, needs of stakeholders, measurability, and cost.

Utility

Before selecting an assessment method to evaluate a library instruction program, teaching librarians should determine the utility of the method. Bresciani suggests asking three questions to determine the utility of a measure: "Will this assessment method help me understand what it is that is contributing to the end result stated in [the measured] outcome? Will this assessment method help me understand why I am delivering the services in the way that I am? Will the evidence collected from this method help me understand how to maintain status quo?"²⁹

Relevance

In addition to utility, librarians should consider the relevance a particular assessment method has to learning and ask themselves whether or not the assessment method will answer questions about student learning. Grassian and Kaplowitz confirm that, "assessment methods should be selected because of their relevance to the learning outcomes or performances to be measured."³⁰ Prus and Johnson state unequivocally that, "If an assessment method doesn't measure what your program teaches, or doesn't measure it with precision, or doesn't suggest what the program's strengths and weaknesses are, then that assessment method cannot serve the institutional effectiveness goals of your program."³¹

Stakeholder Needs

When selecting an assessment tool, librarians should consider the needs of stakeholders, including both the audience that assessment data will be reported to and the participants who will undergo assessment. Stakeholders need assessment data to be displayed in a format that is easily understood.³² Librarians should consider who will see final results and use knowledge of stakeholders when determining the level of preciseness, level of detail, and speed of reporting.³³ Finally, librarians undertaking assessment should reflect upon the needs of those who are assessed, including how many there are, who they are, and what their assessment experience might be.³⁴

Measurability

Two additional considerations in choosing a method for assessment are measurability and cost. Colton, et al. defines measurability as whether or not the tool measures what it intends to measure (validity) with consistency (reliability), and he lists measurability as the first focus in choosing an assessment tool.³⁵ Prus and Johnson define internal validity as "maximum relevance to the unique aspects of the local program curriculum" and external validity as "maximum generalizability to similar programs at colleges across the state, region, and nation."³⁶ These aspects of measurability can significantly impact the credibility of an assessment method.³⁷

Cost

A final key consideration in selecting an assessment method is cost.³⁸ Ideally, educators should choose the assessment method that best fits their programmatic and instructional needs, but cost is a practical concern that cannot be ignored. Cost is impacted first by the scope of an assessment and whether entire populations must be assessed or a sample will suffice. Time is another "cost" that should be considered. In fact, time is one of the most common reasons for not conducting assessment at all.³⁹ Librarians should consider the time that is required for creating a tool, implementing it and analyzing it.⁴⁰ Prus and Johnson agree that librarians should consider whether the cost of a particular assessment method is practical, including time, effort, and money in their deliberations.⁴¹

Outcomes-Based Assessment Tools

Once librarians articulate their purposes for

assessment and criteria for selecting an assessment tool, the next step is to select an assessment approach. Fortunately, there are a wide variety of assessment tools from which librarians can choose including surveys, standardized tests, performance assessments, and rubrics.

Surveys and Standardized Tests

In academic libraries, outcomes-based surveys "provide information about the students' library skills before [or] after a sequence of library instruction and/or research activities."⁴² Some libraries have endeavored to transform outcomes-based surveys into standardized tests in order to administer and score the assessments in a standard, predetermined way⁴³ and to strive for objectivity.⁴⁴ Most outcomes-based surveys and standardized tests focus on multiple-choice or true/false items.⁴⁵

Outcomes-based surveys and standardized tests offer a number of benefits to librarians seeking to engage in assessment. Indeed, librarians have "realized some success with summative assessment devices, such as tests and surveys."⁴⁶ As quantitative measures, surveys and tests provide data in numerical form and are excellent choices to find answers to questions of how much or how many.⁴⁷ They are easy to score and require less time and money, especially if computers are used for scoring.⁴⁸ In this way, they allow for the collection of a lot of data quickly.⁴⁹ Surveys and tests are good tools for measuring students' acquisition of facts⁵⁰ and can be used to compare pre- and post-test results⁵¹ or to compare groups of students to each other.⁵²

Another advantage of surveys and tests, especially those made up of multiple-choice items, is that they can be made highly reliable.⁵³ In fact, high reliability is one of the most frequently cited advantages of surveys and tests.⁵⁴ One way to increase the reliability of a multiple-choice survey or test is to make it longer.⁵⁵ Lengthening a survey or test is much easier than lengthening other types of assessment methods.⁵⁶ Furthermore, test/retest and parallel forms reliability estimates are easier to obtain with surveys and tests than other assessment methods that take more time or are difficult to repeat exactly.⁵⁷ Indirect assessments like surveys and tests also tend to have a higher predictive validity with "a variety of outcome measures, such as college GPA or scores on other standardized tests."⁵⁸

Still another advantage to using outcomes-based surveys and tests for assessment is that people

believe in them.⁵⁹ Because the public is familiar with commercially designed tests and believes them to be extensively developed, tests can be used for "enhanced political leverage."⁶⁰ Policy makers may prefer standardized tests because they compare students' achievement against other groups or national profiles.⁶¹ Parents and students might also value such normative comparisons and try to use them to identify an individual student's strengths and weaknesses.⁶²

Locally developed outcomes-based surveys have several additional benefits. First, they have the benefit of being adapted to local goals and student characteristics.⁶³ The process of developing the surveys can help staff determine what they really want to know about student learning.⁶⁴ Local grading is an additional benefit—staff have control over the interpretation and use of the results and students receive immediate feedback.⁶⁵ Commercially developed standardized tests also offer two benefits: they can be implemented quickly and they reduce the staff time that would be otherwise used to develop and grade another assessment measure.⁶⁶

On the other hand, outcomes-based surveys and tests have many limitations, and while outcomes-based surveys and standardized tests are widely used, most educators now recognize that they "have always been fallible, limited measures of learning goals."⁶⁷ Overall, outcomes-based surveys and standardized tests have several limitations: they do not test higher-level thinking skills, they lack authenticity, they focus on "score-spread," they tend to have the problems associated with all high-stakes testing, and they are time consuming to create, difficult to analyze, and problematic on a local level.

A major limitation of outcomes-based surveys and standardized tests is that they are indirect assessments that fail to measure higher-order thinking skills.⁶⁸ As "objective" tests, they measure low-level recognition rather than recall.⁶⁹ Because of artificial time limits and the pressure to survey as much content as possible, outcomes-based surveys and tests rarely involve the interrelation of dimensions of the same topic.⁷⁰ By focusing only on individual parts of a concept, test creators tend to develop over-simplified test items. Because of this limitation, it is very difficult to use outcomes-based surveys and standardized tests to quantitatively measure the results of improved information literacy instruction.⁷¹ In fact, indirect assessments, like outcomes-based surveys and standardized

tests, may "dramatically under-represent" constructs like writing, critical thinking, and information literacy.⁷² Grassian and 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 [librarians'] needs."⁷³

A second limitation of outcomes-based surveys and standardized tests is their inability to provide an authentic assessment of student learning. These types of tests are "frequently criticized for setting up an artificial situation that does not really test how the learner would react in a real-world situation."⁷⁴ Because of this, outcomes-based surveys and standardized tests tend to "overassess student 'knowledge' and underassess student 'know-how with knowledge.'"⁷⁵ As a result, students who score well on outcomes-based surveys and standardized tests may only be demonstrating that they are good test takers.⁷⁶ When faced with a real-world scenario, these students may not be able to formulate an appropriate response.⁷⁷

A third important limitation of outcomes-based surveys and standardized tests is that most are designed to produce variance of scores, or "score-spread." Most standardized tests are intended to allow comparisons among students or groups of students. To do this, the tests must spread out student scores, rather than allow them to bunch together. The time constraints of most standardized tests exacerbate the problem of score-spread because test designers strive for maximum score-spread in the fewest number of test items.⁷⁸ The use of score-spread techniques results in the creation of tests that cannot detect effective instruction⁷⁹ and, according to Popham, are inappropriate for assessing program or institutional effectiveness.⁸⁰ Even tests that are purported to be criterion-referenced, rather than norm-referenced, are usually constructed in the same way as traditional standardized tests, with the same focus on score-spread. Therefore, they are still "unsuitable-for evaluating educational quality."⁸¹

A fourth significant limitation of outcomes-based surveys and standardized tests of information literacy instruction is that they are sometimes used as "high-stakes" tests. Shepard explains:

High-stakes testing is a term that was first used in the 1980s to describe testing programs that have serious consequences for students or

educators. Tests are high-stakes if their outcomes determine such important things as promotion to the next grade, graduation, merit pay for teachers, or school rankings reported in a newspaper. When test results have serious consequences, the requirements for evidence of test validity are correspondingly higher.⁸²

Because of the high stakes of such assessments, tests that fall into this category must "meet the most stringent technical standards because of the harm to individuals that would be caused by test inaccuracies."⁸³ 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.⁸⁴ Prus and Johnson also caution that standardized test results are "highly susceptible to misinterpretation and misuse both within and outside the institution" and are "unlikely to have direct implications for program improvement or individual student progress."⁸⁵

Three more limitations round out the problems associated with outcomes-based surveys and standardized tests. Surveys and tests are time-consuming, demand significant resources, and provide inferential results.⁸⁶ Grassian and Kaplowitz state that "fixed choice assessments (multiple-choice, true/false, matching) are high on control but are difficult and time consuming to construct. They require a good deal of specialized training to develop and analyze."⁸⁷ If test data is not analyzed or interpreted, such assessments drain resources. Even so, Ewell and Jones suggest that "conclusions drawn from indirect indicators [such as surveys and standardized tests] are highly inferential even when the data are presented from multiple measures."⁸⁸

When outcomes-based surveys and standardized tests are locally developed, a few additional limitations apply. The process of constructing a survey is a difficult one. Locally developed surveys and tests require leadership, coordination, and expertise in measurement.⁸⁹ A survey with good psychometric properties can take years to develop,⁹⁰ and an outcomes-based survey or test may take even longer since they require not just content expertise, but also expertise in the study of learning.⁹¹ As a result, locally-developed surveys and tests may not provide for "externality" or a "degree of objectivity associated with review and comparisons external to the program or institution."⁹² Even if a local assessment is

determined to be adequate, it's important not to administer the same test so often that students become "over-surveyed."⁹³

Performance Assessments

Many modern educators feel a "growing dissatisfaction with selected-response testing."⁹⁴ Because of the limitations of outcomes-based surveys and standardized tests, the emphasis on quantitative results that once dominated assessment conversations in higher education is beginning to give way to discussions of qualitative forms of assessment.⁹⁵

There are a number of ways to structure performance assessment. Performance assessments generally "simulate as much as possible the situations in which students would make integrated use of the knowledge, skills, and values developed in the course."⁹⁶ For example, instead of scoring a student's answers to a set of answers provided by a teacher,⁹⁷ a performance assessor might observe a student's performance of a task or a product of a performance, and judge its quality.⁹⁸ Other methods that have been used for performance assessments are open-ended or extended response exercises (such as questions or other prompts that ask students to explore a topic in writing), extended tasks, or portfolios.⁹⁹ Constructed-response tasks have become popular because they capture what is valued instructionally in a form that can be easily assessed.¹⁰⁰ In general, "Most educators regard performance assessment as an attempt to measure a student's mastery of a high-level, sometimes quite sophisticated skill through the use of fairly elaborate constructed-response items and a rubric."¹⁰¹

Performance assessments should exhibit a number of characteristics. Performance assessments should be meaningful and authentic.¹⁰² Wiggins states that performance assessments should involve actual "performances, not drills. A test of many items (a drill) is not a test of knowledge in use. 'Performance' is not just doing simplistic tasks that cue us for the desired bit of knowledge. It entails 'putting it all together' with good judgment; good judgment cannot be tested through isolated, pat drills."¹⁰³ Shepard agrees that assessments should "require more 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."¹⁰⁴ Shepard also notes that performance assessments should be

open-ended enough to allow each student "to bring to it his individual gifts and to maximize individual learning."¹⁰⁵

Performance assessment offers numerous benefits to teachers and learners. Among them are the close connections between instruction and assessment, the ability to measure higher-order thinking skills, and the contextualization of assessment that leads to greater equitability and validity.

Perhaps the greatest value of performance assessment is that the form and content of the assessment method can be closely aligned with instructional goals.¹⁰⁶ 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."¹⁰⁷ Because instruction and assessment are integrated in performance assessments, educators can learn about a broader range of learning outcomes and, at the same time, "preserve the complex nature of disciplinary knowledge and inquiry, including conceptual understanding, problem-solving skills, and the application of knowledge and understanding to unique situations."¹⁰⁸ Furthermore, performance assessments allow educators to capture students' learning of higher-order thinking and reasoning skills, skills that are typically absent in more traditional forms of assessment.¹⁰⁹

Another benefit of outcomes-based performance assessments is that they are contextualized. According to 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."¹¹⁰ 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."¹¹¹ Through contextualization, performance assessments help students understand the relevance of what they learn.¹¹² Performance assessments also can "reflect . . . society's demands for education that prepares students for the world of work."¹¹³ According to Farmer, "Authentic assessment helps bridge the two worlds. . . . They show that [students] can apply theoretical concepts to solving life-like problems. As a result, education doesn't seem to operate in a vacuum; it truly prepares students for the rest of their lives. [It is] real learning for real results."¹¹⁴

According to Silver, there are a few technical and feasibility issues that have, in the past, thwarted attempts to use performance assessment on a large scale.¹¹⁵ Among these are cost, time, and generalizability. For example, compared to traditional tests, performance assessments can be costly to develop, administer, and score.¹¹⁶ However, Silver offers hope that, in the future, new technologies will decrease the costs associated with test development, administration, and scoring.¹¹⁷ Another limitation of an outcomes-based performance assessment approach is time. Sweet notes that performance assessments require greater time, in both planning and thought, from both teachers and students.¹¹⁸

A third limitation of outcomes-based performance assessment approaches is the generalizability and comparability of results.¹¹⁹ Popham asks, "How many performance tests do students need to complete before the teacher can come up with valid inferences about their generalizable skill-mastery?"¹²⁰ He fears that many tests may be needed and, as a result, cautions educators to use performance assessment to assess "only the most significant of your high-priority curricular aims."¹²¹ Still, this concern about the generalizability of performance assessments may be only temporary. According to Silver, "Generalizability across tasks may be increased through the use of intelligent systems that offer ongoing assessment well integrated with instruction and sensitive to changes in students' understanding and performance, with performance data collected over a long period of time as opposed to one-time, on-demand testing."¹²² Overall, Silver suggests that advances in cognitive sciences and technology may resolve the limitations of performance assessment.¹²³

Rubrics

Rubrics are "descriptive scoring schemes" created by educators to guide analysis of student work¹²⁴ that are often employed when educators must judge the quality of performances or constructed-response items.¹²⁵ In fact, rubrics are ubiquitous in educational assessment and offer substantial benefits to librarians conducting assessment of information literacy skills and library instruction.

The benefits of rubric assessment are numerous.¹²⁶ Rubrics have instructional advantages over other assessment approaches, including the direct benefits to students, the benefits of stating agreed upon values, and the benefits of detailed

result data. Popham states, "A properly fashioned rubric can help teachers teach much more effectively and help students learn much more effectively, too."¹²⁷ Pausch and 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."¹²⁸ Most importantly, rubrics benefit students in several ways. First, rubrics allow students to understand the expectations of their instructors. Second, rubrics provide direct feedback to students about what they have learned and what they have yet to learn. They also can facilitate student self-evaluation. Finally, rubrics are used to make rankings, ratings, and grades more meaningful by revealing what educators expect students to know or do.¹²⁹

Rubric assessment provides another important instructional benefit—the opportunity to discuss and determine agreed upon values of student learning. Callison writes, "Rubrics are texts that are visible signs of agreed-upon values. They cannot contain all the nuances of the evaluation community's values, but they do contain the central expressions of those values."¹³⁰ Stevens and Levi list the facilitation of communication with others as a key reason to use rubrics.¹³¹ Bresciani, Zelna and Anderson confirm that rubrics "make public key criteria that students can use in developing, revising, and judging their own work."¹³² They also point out that once rubrics are developed, they can be used to norm educators' expectations and to bring them in line with the original vision for student learning.¹³³ They can also be used to make students full participants in the assessment process. Rubrics allow the possibility of including students in the development of standards for a product or performance.¹³⁴ Bernier points out that "helping teachers create rubrics helps students meet those content standards."

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."¹³⁵ This descriptive data can be used to document how educators or program administrators improve instruction.¹³⁶ Furthermore, the data that results from rubric assessment is so detailed and well-defined that it "combats accusations that evaluators do not know what they are looking for in learning and development."¹³⁷ The level of detail found in rubrics helps prevent inaccuracy of scoring¹³⁸ and

bias.¹³⁹ Rubrics clarify schemes for assessment ahead of time, and therefore reduce subjectivity in grading.¹⁴⁰ Since rubrics guide teachers to focus on essential criteria,¹⁴¹ they can grade student products and performances more easily and objectively.¹⁴² Callison indicates the importance of this benefit for librarians, stating that rubric assessment "is more likely to be reasonably objective and consistent from lesson to lesson and from student to student, especially useful in team teaching situations that involve collaboration among library media specialists and other teachers."¹⁴³ Because rubrics are easy to use and easy to explain, they also generate data that is easy to understand, defend, and convey.¹⁴⁴

Like other assessment tools, 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,¹⁴⁵ and crafting a good rubric requires time, practice, and revision.¹⁴⁶ Tierney and 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."¹⁴⁷

Another limitation of rubric assessment is time. While creating rubrics is inexpensive monetarily, some teachers find the process time-consuming.¹⁴⁸ Part of that perception might be due to lack of familiarity or expertise; teachers don't always know how to make a rubric and so they believe the process will take too much time.¹⁴⁹ Prus and Johnson acknowledge the potential cost of time required to create a rubric, but feel that the advantages outweigh the costs.¹⁵⁰ They 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 faculty time, student effort, and money."¹⁵¹ Stevens and Levi argue that rubrics actually make grading easier and faster by "establishing performance anchors, providing detailed, formative feedback, . . . supporting individualized, flexible, formative feedback, . . . and conveying summative feedback."¹⁵²

Conclusion

Although selection of an assessment method may seem a daunting task, it is one that can be guided by criteria for selection and best practices. Librarians attempting to select assessment methods

must understand the strengths and weaknesses of various techniques.¹⁵³ They also must consider the types of information that various techniques provide, and whether or not the methods they select will answer the questions they have about their library instruction programs.¹⁵⁴ Ultimately, the choice of assessment comes down to a fit between the purposes of assessment and the capabilities of assessment methods. As Grassian and Kaplowitz point out, "An assessment method is neither good nor bad in its own right."¹⁵⁵ Success in assessment is simply a matter of selecting the right tool for the job.

—Copyright 2007 Megan J. Oakleaf

Endnotes

1. Lorrie A. Shepard, "Standardized Tests and High-Stakes Assessment," in *Encyclopedia of Education*, ed. James W. Guthrie, vol. 7 (New York: Macmillan Reference USA, 2003).
2. Lois M. Pausch and Mary Pagliero Popp. "Assessment of Information Literacy: Lessons from the Higher Education Assessment Movement," ACRL National Conference (Nashville, TN: Association of College and Research Libraries, 1997).
3. Bonnie Gratch Lindauer, "Comparing the Regional Accreditation Standards: Outcomes Assessment and Other Trends," *Journal of Academic Librarianship* 28, no. 1 (2002): 16.
4. Ralph A. Wolff, "Using the Accreditation Process to Transform the Mission of the Library," *New Directions for Higher Education* 90 (1995): 80.
5. Gary B. Thompson, "Information Literacy Accreditation Mandates: What They Mean for Faculty and Librarians," *Library Trends* 51, no. 2 (2002): 218.
6. Bonnie Gratch Lindauer, "Defining and Measuring the Library's Impact on Campuswide Outcomes," *College and Research Libraries* 59, no. 6 (1998): 549.
7. *Ibid.*, 546-547.
8. Gratch Lindauer, "Comparing the Regional Accreditation Standards," 20.
9. Edward K. Owusu-Ansah, "Information Literacy and Higher Education: Placing the Academic Library in the Center of a Comprehensive Solution," *Journal of Academic Librarianship* 30, no. 1 (2004): 6.
10. Marilee Bresciani, Carrie Zelna and James Anderson, *Assessing Student Learning and Development*, (National Association of Student Personnel Administrators, 2004), 16.
11. *Ibid.*, 16.
12. Sue Samson, "What and When Do They Know? Web-Based Assessment," *Reference Services Review* 28, no. 4 (2000): 335-342.
13. Lorrie A. Knight, "The Role of Assessment in Library User Education," *Reference Services Review* 30, no. 1 (2002): 15-24.
14. Donald Barclay, "Evaluating Library Instruction: Doing the Best You Can with What You Have," *Research Quarterly* 33, no. 2 (1993): 195-203.
15. Bresciani, Zelna and Anderson, 16.
16. *Ibid.*, 17.
17. Knight.
18. Esther S. Grassian and Joan R. Kaplowitz, *Information Literacy Instruction*, (New York: Neal-Schuman, 2001), 266.
19. *Ibid.*, 266.
20. Knight.
21. James W. Popham, *Test Better, Teach Better* (Alexandria, Virginia: Association for Supervision and Curriculum Development, 2003), 69.
22. Research and Scholarship Committee of the ACRL Instruction Section, "Research Agenda for Library Instruction and Information

- Literacy," *College and Research Libraries* 64, no. 2 (2003): 112.
23. Pausch and Popp.
24. Grassian and Kaplowitz, 265.
25. Samson.
26. Judith A. Arter, "Using Assessment as a Tool for Learning," in *A Handbook for Student Performance in an Era of Restructuring*, eds. Robert E. Blum and Judith A. Arter (Alexandria, VA: Association for Supervision and Curriculum Development, 1996), IV-10:1.
27. Peggy L. Maki, "From Standardized Tests to Alternative Methods: Assessing Learning in Education," *Change* 33, no. 2 (2001): 15-24.
28. Bresciani, Zelna and Anderson, 25.
29. *Ibid.*, 20.
30. Grassian and Kaplowitz, 9.
31. Joseph Prus and Reid Johnson, "A Critical Review of Student Assessment Options," *New Directions for Community Colleges* 88 (1994): 70.
32. 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), 2-3.
33. Grassian and Kaplowitz, 272.
34. Dean A. Colton et al., *Reliability Issues with Performance Assessments: A Collection of Papers* (Iowa City: ACT, 1997), 9.
35. *Ibid.*, 6.
36. Prus and Johnson, 70.
37. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 2-3.
38. Assessment Methods, (Assessment Strategies, Inc., 1999), <http://www.asinc.ca/en/methods.html>.
39. Bresciani, Zelna and Anderson, 26.
40. *Ibid.*, 26.
41. Prus and Johnson, 70.
42. Knight.
43. Popham, 125.
44. Johanna V. Crighton, "Standardized Tests and Educational Policy," in *Encyclopedia of Education*, ed. James W. Guthrie, vol. 7 (New York: Macmillan Reference USA, 2003), 2530.
45. Crighton, 2530.
46. Knight.
47. Grassian and Kaplowitz, 276.
48. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 4.
49. Grassian and Kaplowitz, 281.
50. *Ibid.*, 281.
51. Knight.
52. Grassian and Kaplowitz, 281.
53. Colton et al., 3.
54. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 4.
55. Colton et al., 3.

-
56. Ibid., 3.
57. Ibid., 4.
58. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 4.
59. Lorrie A. Shepard, "Why We Need Better Assessments," in *A Handbook for Student Performance Assessment in an Era of Restructuring*, eds. Robert E. Blum and Judith A. Arter (Alexandria, VA: Association for Supervision and Curriculum Development, 1996), I-2: 3.
60. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 4.
61. Maki.
62. Popham, 126.
63. Prus and Johnson, 72.
64. Ibid., 71.
65. Ibid., 72.
66. Ibid., 71.
67. Shepard, "Why We Need Better Assessments," I-2:6.
68. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 4.
69. Grassian and Kaplowitz, 281.
70. Shepard, "Why We Need Better Assessments," I-2: 3.
71. Molly R. Flaspohler, "Information Literacy Program Assessment: One Small College Takes the Big Plunge," *Reference Services Review* 31, no. 2 (2003): 139.
72. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 4.
73. Grassian and Kaplowitz, 281.
74. Ibid., 281.
75. Grant Wiggins, "Creating Tests Worth Taking," in *A Handbook for Student Performance in an Era of Restructuring*, eds. Robert E. Blum and Judith A. Arter (Alexandria, VA: Association for Supervision and Curriculum Development, 1996), V-6: 3.
76. Grassian and Kaplowitz, 277.
77. Ibid., 277.
78. Popham, 127.
79. Ibid., 128.
80. Ibid., 127.
81. Ibid., 140.
82. Shepard, "Standardized Tests and High-Stakes Assessment," 2533-2534.
83. Ibid., 2535.
84. Prus and Johnson, 71-72.
85. Ibid., 71-72.
86. National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 4.
87. Grassian and Kaplowitz, 281.
88. Ewell and Jones in National Postsecondary Education Cooperative Student Outcomes Pilot Working Group: Cognitive and Intellectual Development, 4.
89. Prus and Johnson, 72-73.

90. Bresciani, Zelna and Anderson, 70.
91. Kenneth R. Smith, *New Roles and Responsibilities for the University Library* (Association of Research Libraries, 2000).
92. Prus and Johnson, 72-73.
93. Bresciani, Zelna and Anderson, 71.
94. Edward A. Silver, "Performance Assessment," in *Encyclopedia of Education*, ed. James W. Guthrie, vol. 1 (New York: Macmillan Reference USA, 2003), 135.
95. Pausch and Popp.
96. David Sweet, "Performance Assessment," *Education Research Consumer Guide 2* (1993).
97. Ibid.
98. Silver, 134.
99. Sweet.
100. Silver, 135.
101. Popham, 101.
102. Wiggins, V-6: 3.
103. Ibid., V-6: 4.
104. Shepard, "Why We Need Better Assessments." I-2: 4.
105. Ibid.
106. Silver, 135.
107. Ibid., 135.
108. Ibid., 135.
109. Ibid., 135.
110. Wiggins, V-6: 5.
111. Ibid., V-6: 3.
112. Lesley s. Johnson Farmer, "Authentic Assessment of Information Literacy through Electronic Products," *Book Report 16* (1997): 11-13.
113. Ibid.
114. Ibid.
115. Silver, 135.
116. Ibid., 135.
117. Ibid., 136.
118. Sweet.
119. Silver, 135.
120. Popham, 102.
121. Ibid., 103.
122. Silver, 136.
123. Ibid., 136.
124. Barbara M. Moskal, "Scoring Rubrics: What, When, and How?" *Practical Assessment, Research, and Evaluation 7*, no. 3 (2000).
125. Popham, 95.
126. Bresciani, Zelna and Anderson, 30.
127. Popham, 95.
128. Pausch and Popp.
129. Bresciani, Zelna and Anderson, 31.
130. Daniel Callison, "Rubrics," *School Library Media Activities Monthly 17*, no. 2 (2000): 36.
131. 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), 23.

-
132. Bresciani, Zelna and Anderson, 30.
133. Ibid., 31.
134. Ibid., 31.
135. Ibid., 30.
136. Rosemarie Bernier, "Making Yourself Indispensible by Helping Teachers Create Rubrics," *CSLA Journal* 27, no. 2 (2004): 25.
137. Bresciani, Zelna and Anderson, 30.
138. Popham, 95.
139. Bresciani, Zelna and Anderson, 31.
140. Moskal.
141. Callison, 35.
142. Bernier, 25.
143. Callison, 35.
144. Heidi Goodrich Andrade, "Using Rubrics to Promote Thinking and Learning," *Educational Leadership* 57, no. 5 (2000): 14.
145. Popham, 95.
146. Callison, 35.
147. Robin Tierney and Marielle Simon, "What's Still Wrong with Rubrics: Focusing on the Consistency of Performance Criteria across Scale Levels," *Practical Assessment, Research, and Evaluation* 9, no. 2 (2004).
148. Ibid.
149. Bernier, 25.
150. Prus and Johnson, 25.
151. Ibid., 25.
152. Stevens and Levi, 73.
153. Grassian and Kaplowitz, 279.
154. Ibid., 273.
155. Ibid.