Use What You Have: Authentic Assessment of In-Class Activities

Introduction

Not exempt from the mounting call in academia for evidence of student learning, libraries must prove their value on the educational experience. The voluminous literature on the assessment of information literacy indicates that libraries have both recognized these stakes and have taken action. Case studies and research papers abound on any number of methods used alone or in combination to evaluate libraries’ effectiveness in information literacy instruction. For example, assessment may occur at the institutional, program, course, or session levels (Gilchrist and Oakleaf, 2012). It may take the form of traditional, authentic, quantitative, qualitative, formative, summative, formal, informal – the list goes on. Regardless of the level or method chosen, one objective should serve as the foundation for all assessments – the goal of “closing the loop” (Maki, 2004) of the Information Literacy Instruction Assessment Cycle (ILIAC). Outlined by Oakleaf (2009), this loop assists in determining if students meet learning goals and provides evidence on what changes should be made to improve the results.
For Auburn University’s library instruction program, the aim of closing the loop has taken center-stage within the past few years, resulting in the adoption and use of an array of assessment methods. Nevertheless, the identification of a gap in these practices – that of measuring student learning of outcomes within individual sessions of course-integrated information literacy – led to the development of a formative assessment technique that provides immediate feedback to librarians regarding their teaching strategies. The implementation of this piece of assessment consists of three key elements: 1) The use of authentic assessment offers librarians an analysis of how students apply the outcomes taught in a specific class. Did they choose the appropriate keywords to search for their topics? Did they properly identify scholarly sources for their papers? Did they successfully determine the reliability of websites?; 2) Because all librarians at Auburn University Libraries integrate at least one active learning component into each class, the authentic assessment of that component results in a seamless process. This solves the challenge of devoting limited class time to assessment while still sufficiently addressing the learning outcomes.; 3) The decision to concentrate on outcomes within the ACRL’s information competency standards (Association of College and Research Libraries, 2000), rather than on a broader scale (such as measuring student learning standards or performance indicators) provides a focused assessment. Librarians can now point to minor, yet critical modifications they can make to their teaching that will enhance student learning. This paper explains the methodology used in authentically assessing students’ abilities to “identify keywords, synonyms, and related terms for the information needed” (from ACRL outcome 2.2.b.). It will also include a discussion of how students performed on the assessments and what librarians will adjust in order to close the loop.
Background and Context

Instruction program

Auburn University Libraries (AUL) serves approximately 25,000 students. A robust and aggressive instruction program plays a significant role in this service, with seventeen teaching librarians conducting around 600 to 700 information literacy sessions per year. The ACRL’s *Information Literacy Competency Standards for Higher Education* (Association of College and Research Libraries, 2000) along with a philosophy of active learning in the classroom function as the backbone of the program. Four of the seventeen teaching librarians serve on an instruction team (led by the instruction coordinator) steering the ship of information literacy by sharing teaching ideas, heading instruction and assessment initiatives, and encouraging librarians to collaborate with faculty across campus. Subject specialists conduct library classes in their disciplines, but most teaching librarians lead library instruction sessions for second-semester freshman composition (ENGL1120). The assessment method discussed in this paper has mainly occurred in these composition classes, although the instruction team actively encourages its use in discipline-specific classes as well.

The freshman composition course represents the most vigorous partnership on campus for the library instruction program, with all composition classes attending three library instruction sessions. The luxury of meeting with the students three times allows librarians to limit the number of outcomes to one or two per class, affording a more comprehensive focus on specific skills than would most likely occur in a one-shot session. Library instruction complements the three papers within the composition curriculum in which instructors require students to use outside sources: comparison and contrast of two sources, syntheses of an argument, and a final research paper. Librarians reach students at their point of need by scheduling the sessions after the class receives each assignment, but before research and writing
begin. Further, these papers build upon each other, and librarians use this same scaffolding approach when defining the learning outcomes for each session. For instance, the first class would introduce competences needed to find sources in a database, such as identifying keywords and synonyms for the students’ topics of choice, followed by an introduction to Academic Search Premier. Composition instructors require scholarly articles for the next paper, therefore evaluating sources would most likely be appropriate for the second class. The third session could focus on subject databases and using the link resolver to locate full-text. The outcomes may vary slightly from these examples, and librarians negotiate with the instructors in regard to what skills the students need in order to complete their assignments. For additional information on the creation of this curriculum and the collaboration between AUL and the composition program, see Rumble and Noe (2009).

Assessment

Gilchrist and Oakleaf (2012), prolific writers on the topic of information literacy assessment, define four levels of outcomes assessment – institution, program, course and individual session. AUL participates in each of these except course-level due to the absence of an information literacy credit course. In 2003, with a Southern Association of Colleges and Schools (SACS) accreditation on the horizon, programmatic assessment began with AUL’s involvement in the pilot testing of Project SAILS (Standardized Assessment of Information Literacy Skills) (Project SAILS, 2000-2012) (Rumble and Noe, 2009). Since then, this quantitative exam has been administered in alternating years, providing AUL with a measure of the information literacy skills of Auburn students compared to students from across the country. AUL’s work with Project SAILS led to the adoption in 2008 of information literacy (based on the ACRL information competency standards) as a student learning outcome within the
university’s core curriculum. Thus, Project SAILS now operates as both a programmatic and institutional-level assessment, as defined by Gilchrist and Oakleaf (2012). The instruction team also conducts summative assessment of course-integrated information literacy through citation analyses of students’ final research papers for ENGL1120 and Professional Concepts I (Nursing 3110).

Although Project SAILS assists AUL in proving the value of its services to the University, and citation analyses offers a measure of whether students can apply multiple skills taught in library instruction classes to their final papers, both approaches lag in providing guidance on specific teaching adjustments that could improve student learning. They inform librarians where the students excel or where they struggle, but do not necessarily explain why they perform in a certain manner. To gather more granular feedback, librarians have recently explored and implemented two formative types of assessment for use in individual instruction sessions. First, through the use of Classroom Assessment Techniques (CATs) made popular by Angelo and Cross (1993), librarians administer quick, informal assessments that can be used to gather feedback. Examples of CATs include one-minute papers, muddiest points, and concept maps. However, while the CATs used at AUL encourage students to ask questions and describe what they have learned, the second method adopted by the instruction program – authentic assessment – compels the students to show what they have learned. While assessment at the individual session-level delivers immediate feedback, librarians expect the integration of authentic assessment as a standard procedure to result in longitudinal improvements at the programmatic and institutional-levels as well.
**Literature Review**

A review of the literature uncovers a plethora of sources that examine the definition and benefits of authentic assessment. While most studies point to agreement concerning the beneficial nature of this practice, the definition of authentic assessment offers fodder for debate. Most experts concur, though, that the core of authentic assessment involves challenging students with real-world/real-life scenarios. Rather than adding to that considerable discussion, however, this review focuses on putting authentic assessment into practice. Even this concentrated emphasis reveals a wide-variety of methods. Narrowing down to the measurement of student learning outcomes within individual sessions of course-integrated information literacy confirms that summative assessment dominates the literature. In fact, studies on the equally essential formative assessment of student work, particularly work completed in individual instruction sessions, prove scant in the library literature. However, examples of formative assessments outside of the library literature – beyond the scope of this review but worth mentioning – may provide transferrable ideas to the information literacy realm. For instance, Brookhart (2012) discusses teacher feedback to students when using formative assessments, while Hessler and Taggart (2011) explain their use of Brookfield’s (1995) “Critical Incident Questionnaire” (although this aligns more closely with the CATs (Angelo and Cross (1993)) employed at AUL than the authentic assessments of real-world situations). Instructors of biology classes at the University of Sydney experimented with non-required formative assessments such as practice exams and self-assessments, with the goal of improving student performance on later summative assessments (Peat *et. al*, 2004). In the absence of an exact model of AUL’s assessments, the most comparable methods within library literature will be shared, including assessments measuring on a broader scale than outcomes. This sampling of the literature verifies that librarians view authentic assessment as a viable way to close the loop.
Outcomes-based

In one of the more parallel methods to AUL’s, Chen (2009) describes a practice of measuring students’ success in identifying keywords for their research (ACRL outcome 2.2.b). By asking students several times throughout the research process to amend their keywords, librarians could formatively assess if improvements to keyword selection occurred after instruction. This assessment took place in a semester-long information literacy class, defined by Gilchrist and Oakleaf (2012) as course-level. In the absence of such a course, librarians at AUL must evaluate students’ identification of keywords at the individual session-level. In another semester-long course, SUNY-Brockport (Nutefall, 2005) introduced an innovative assessment utilized in an oral communication class co-taught by the Communication Department and the Library. Called the “Paper Trail,” the assessment prompted students to describe aspects of their research, with some questions focusing on specific outcomes. Examples include “what search terms do you use?” and “did you try different terms?” By requiring students to turn in the “Paper Trail” before the end of the semester, time remained for librarians to work on problem areas of student learning.

Returning to course-integrated rather than course-level assessment, librarians at Hong Kong Baptist University (Cmor et al., 2010) examined in-class activities (similar to AUL) to determine if students successfully developed database searching strategies. While this “outcome” is technically a performance indicator according to the ACRL standards, no mention of ACRL appears in the paper, possibly due to differences in U.S. and non-U.S. libraries. Librarians did, however, separate out the identification of keywords from other components of database searching strategies, such as the use of Boolean. Turning to a different outcome,
Daniels (2009) reports on an assessment of students’ abilities to evaluate the credibility of sources. Determining that programmatic assessment of performance indicators failed to measure exactly where in the process of evaluating sources students struggled, librarians at Sonoma State University developed a “targeted” rubric to assess the “smaller learning stages required for a student to successfully evaluate a source.” Although summative, the author maintains that the process helped librarians pinpoint in which “developmental stages” the students lagged. This work of recognizing the “smaller learning stages” that influence student performance plays a vital role in AUL’s efforts. Lastly, two recent presentations at professional conferences address formative, outcomes-based assessment in one-shot instruction sessions (Acosta, 2012) (Willson, 2011). Hopefully these presenters will share their findings in writing.

Performance indicators and standards-based

Emmett and Emde (2007) report on the use of summative assessment of standards, performance indicators, and outcomes in the form of pre- and post-tests. While typically considered traditional assessment, librarians at the University of Kansas cleverly developed pre- and post-tests of authentic “information-need scenarios” for students to solve. In an analogous example of the usage of pre- and post-tests and real-life scenarios, Brown and Kingsley-Wilson (2012) report on the mapping of ACRL standards to the outcomes of a journalism course. They aimed to measure “if their students were learning the skills they were trying to teach.” The authors coined the phrase “organic assessment” to describe the process of turning an assignment, in this case a journalism assignment, into an authentic assessment tool. Flaspholer (2003) discusses the use of active learning techniques to reinforce skills taught in class (similar to AUL), although it appears librarians assessed standards, not outcomes, and rather than examining the in-class activities, they employed summative assessment of student
bibliographies. Knight (2002), too, describes the use of worksheets in class to measure students’ application of the information taught. However, the librarians were assessing “student competency in library skills,” and the article does not specify whether these included outcomes, performance indicators, or standards. Examples of authentic assessment of standards at the programmatic-level exist as well. For instances of these, see Brown et.al (2003), Diller and Phelps (2008), and Warner (2003).

Methodology

Through their participation in Project SAILS, librarians at AUL obtain data regarding student learning of information literacy at the programmatic and institutional-levels. While this data assists in improving student performance and closing the loop, AUL’s instruction coordinator envisioned an additional assessment method that offered fast and direct feedback of student learning within single classes. Librarians would then have a clearer sense of successful teaching strategies versus techniques that required fine-tuning. Therefore, in the spring of 2008, the coordinator charged one of two newly hired instruction librarians with leading efforts to integrate assessment into individual library instruction classes.

Soon after returning from the assessment track of ACRL’s Immersion Program in fall 2009, the new librarian assigned to assessment shared ideas with the instruction team (IT). The team’s brainstorming sessions resulted in a plan of authentic assessment to complement the more traditional Project SAILS. The elements of this plan included the assessment of worksheets already used in instruction sessions and the evaluation of only one ACRL outcome per class. By authentically assessing these previously integrated worksheets, the process would not require the manipulation of class time to “fit something else in.” And, with each instruction session already limited to one or two outcomes, the measurement of just one would not necessitate major
adjustments to teaching librarians’ lesson designs. The plan also called for the IT to experiment with this approach before introducing it to their colleagues, as well as focusing the initial stages of the project on the instruction program’s bread and butter, freshmen composition (ENGL1120), with the intention of expanding it to upper-level subject classes in the future.

After developing a plan, the next step involved the selection of an outcome to assess. Although not standardized, the ENGL1120 sequence usually covers what the IT terms “the essentials,” with “keywording” at the top of this list. While defined by ACRL’s Information Literacy Competency Standards of Higher Education (Association of College and Research Libraries, 2000) as a “lower order” thinking skill, librarians at AUL consider this outcome a crucial building block for students’ successful performance when applying “higher order” thinking skills. Further, because of numerous definitions given to authentic assessment, some may argue that identifying keywords does not constitute a “real-world” task. The IT deems the ability to properly search for sources as a skill students will not only need in college, but also in the “real-world.” So, the decision to begin with ACRL outcome 2.2.b “identify keywords, synonyms, and related terms for the information needed” (Association of College and Research Libraries, 2000) seemed an obvious choice.

Teaching librarians tackle this outcome in a variety of ways, but most sessions incorporate an introduction to the concept by the use of an example topic or research question. Students are shown how to identify the keywords within the topic, and then how to develop synonyms for the selected keywords. This process often includes the librarian completing a worksheet in front of the class (the same worksheet the students receive) for the example. Students then spend ten to fifteen minutes developing keywords and synonyms for their own
topics with individual worksheets. An introduction to a database, usually Academic Search Premier, concludes the class. Students spend any remaining time searching for sources.

Although librarians at AUL prefer students to enter an instruction session with well-defined topics, this is often not the reality. They therefore consult with composition instructors prior to the sessions, establishing where students situate in topic development for their papers. This plays a critical role at AUL when teaching keywords and synonyms. Some classes may attend an instruction session while still in the early stages of this process, requiring emphasis on keywords, while others may have sufficiently outlined research questions, allowing for more focus on synonyms and related terms. With a multitude of worksheets that cover this variance, librarians employ the one that best meets the needs of the class (for examples, see Figures 1-3) (note: these worksheets serve as templates; students may not need three keywords, or they may identify more than three keywords). Because these in-class activities already existed, the IT concentrated the majority of their efforts on a tool to evaluate the worksheets. After much discussion, the team crafted a rubric that would provide a quantitative and analytic approach by assessing the number of appropriate keywords and synonyms identified by students.

Implementation

With the intention to professionally share preliminary results of student performance on the worksheets, team members needed approval from the Institutional Review Board (IRB). Due to the age of consent in Alabama and IRB rules, only data from worksheets completed by students 19 or older could be included in publications or presentations. The IRB also required consent from the head of the English department. Because the student learning outcomes of the core curriculum place the responsibility of information literacy on ENGL1120, the department
head and director of freshman composition offered full support. Lastly, IRB reviewed all worksheets and rubrics proposed for use. The remaining wrinkle concerned how to collect the worksheets. Librarians needed a copy to review and students needed a copy in order to continue their research after class. While an online form seemed the easiest solution, the IT felt that students would brainstorm more effectively by writing rather than typing. The team also considered photocopying the completed sheets and returning them back to the instructor. In the end, digital carbon paper was the selected medium. It would allow students to retain their worksheets without the need of photocopying, and librarians would walk away with copies as well.

Experimentation with authentic assessment began in the summer and fall of 2010. As the team expected, the number of reviewable worksheets proved low due to fewer ENGL1120 sessions taught during those semesters. Not until the spring, when the majority of these sessions take place, could they accurately measure the success of this assessment. Two issues arose from the initial review. First, although the team had acquired authorization to use several different types of keywording worksheets, they needed more flexibility in developing new ones or tweaking the sheets already accepted by the IRB. Second, the rubric did not address inappropriate keywords and synonyms identified by students, a point that the team felt needed attention. A revised protocol and one-year extension of the project met IRB approval in spring 2011. At the close of the spring 2012 semester, the IT initiated a final review of the process.

Results and Discussion

Evaluation of the process

The team reviewed approximately 250 worksheets over the two-year trial period, enough to provide a suitable evaluation of this process. The use of active learning components as
assessment tools proved successful, most likely a result of their prior integration within instruction sessions. Conversely, the rubrics required more adjustments. After applying the revised rubrics approved by the IRB in spring 2011 to the worksheets, the team realized that the quantitative and analytic-style rubrics failed to meet their needs. First, because the original rubrics seemed too simple, the revised rubrics took into account a wider range of responses, including student selection of inappropriate keywords and synonyms. The results of the assessments revealed the difficulty of this strategy, at least when using these particular worksheets. The team learned they could not systematically measure the variation of student responses with the first two versions of the rubrics. Second, a review of the worksheets prompted a question for the team: what was more important, quantity or quality? Although it may be possible to measure both with just one rubric, the decision by the IT to strip the rubrics of all numbers should result in a qualitative and holistic approach. Librarians will now use the rubrics to evaluate the worksheets as a whole, without calculating the number of appropriate keywords and synonyms.

During this two-year period the team attained IRB approval for assessing a second outcome – identifying scholarly and popular sources. Again, by evaluating active learning modules already utilized in instruction sessions, along with heeding lessons learned from the keywording assessment, implementation of this second outcome has progressed smoothly. The team plans to add more assessments of “the essentials” in the future. Further, with the creation of a “student learning assessment” LibGuide and the hosting of several assessment workshops, the IT has begun sharing ideas with their colleagues. The team encourages teaching librarians to experiment with the keywording and popular/scholarly sources assessments as well as offers guidance to them on how to develop techniques to measure other ACRL outcomes. While it
took the IT two years to “perfect” this process, librarians at AUL should now be on the path to closing the loop at the individual session-level. This authentic method will allow quick and efficient assessment, not only in ENGL1120 sessions, but also in upper-level subject classes.

Results of the assessment

While the purpose of this assessment method aims to provide immediate feedback following individual instruction sessions, throughout the two-year experimental phase the IT gaged student learning of identifying keywords and synonyms. Librarians taught most instances of this outcome during the first instruction sessions of ENGL1120, which usually coincided with the “comparison and contrast of two sources” paper. Results of the worksheets showed that most students successfully identified keywords within their topics or research questions, as well as chose appropriate synonyms and related terms for the keywords. As expected, minor adjustments to teaching techniques should correct weaknesses. These issues will be communicated to other teaching librarians along with suggestions of ways to remedy the concerns and close the loop.

A closer examination of the worksheets reveal that the type of topic (humanities v. science, for instance) mattered little in the quality of the worksheets. However, the IT discovered one matter that involves a separate outcome altogether and possibly skews the positive results found – that of creating realistic and narrow topics. Students easily identified keywords and synonyms from unrealistic or broad topics, but would probably encounter trouble when entering these keywords into a database. The examples below of broad topics with suitable keywords illustrate this disconnect.

Example A

Topic: “who is to blame for childhood obesity?”
As mentioned earlier, librarians ascertain the progress of students in developing their topics by meeting with composition instructors prior to the session. If students have well-developed topics or research questions, librarians may devote more time to teaching how to choose appropriate synonyms. If not, they may allot the majority of class time to identifying keywords. Students’ placement within topic development also determines which worksheet the librarians use. The IT learned, however, that despite the pre-class meetings, either the composition instructors overestimate students’ abilities in achieving this outcome, or librarians fail to properly communicate to the instructors the importance of this step. The team also realized the ineffectiveness of addressing keywords and synonyms when students were still in the “brainstorming” phase. In the examples below, the students clearly chose the wrong keywords. However, the IT postulates that the students also needed more direction in developing their topics. It appears they listed “aspects” of their topics. By focusing on just one of these “aspects,” the students could suitably narrow down their research.

Example A

Topic: “parents are narcissistic towards their children”

Keywords chosen: “sports,” “school,” “life in general”
Example B

Research question: “are transgenic plants good for the environment?”

Keywords chosen: “biopesticide,” “pest control,” “human health issues”

Example C

Topic: “alternative energy options”

Keywords chosen: “wind,” “water,” “solar”

The team is currently devising solutions to address broad topics and/or underdeveloped topics. Improved communication with the composition instructors exemplifies the most obvious answer. In an effort to efficiently address information literacy, these two groups must determine with whom the responsibility for this outcome lies. If with librarians, then an entire class devoted to topic development may prove necessary. On the other hand, if instructors take on this assignment, then librarians must stress the importance of appropriate topics for ensuring a worthwhile instruction session. While members of the IT suspected issues with topic development, without this authentic assessment method they would not have the data needed to close the loop. While challenging, the team views this finding as a positive step forward in improving student learning.

Conclusion

Libraries must employ a variety of assessment methods to accurately measure student learning. They all serve a purpose, and they all have their place. The goal of this case study was to share one practical form of authentic assessment crafted for individual sessions of library instruction. This method does not require an extensive amount of time or work. In fact, the idea of “using what you have” exemplifies this technique. With the popularity of active learning in the classroom, most teaching librarians likely have a module or in-class activity they could easily
transform into an assessment. And, by focusing on the evaluation of ACRL outcomes, the building blocks of performance indicators and standards, librarians can accurately get to the root of teaching and learning issues. In turn, improvements can be made, and librarians will successfully close the loop.

References


Daniels, E. (2010), “Using a targeted rubric to deepen direct assessment of college students’ abilities to evaluate the credibility of sources,” College & Undergraduate Libraries,


