Using Rubrics to Assess Learning in Course-Integrated Library Instruction

Laura W. Gariepy, Jennifer A. Stout, and Megan L. Hodge

abstract: Librarians face numerous challenges when designing effective, sustainable methods for assessing student learning outcomes in one-shot, course-integrated library instruction sessions. We explore the use of rubrics to programatically assess authentic learning exercises completed in one-shot library sessions for a large, required sophomore-level research and writing course. We discuss our findings, as well as how results are used and subsequently impact our information literacy initiatives on a programmatic level. This model is scalable and can be adapted for class size, time, and subject matter.

Introduction

A number of challenges face librarians who seek to assess student learning in course-integrated library instruction sessions. By their nature, one-shot library classes have limited time, requiring librarians to ensure that assessment does not encroach upon the learning itself. In this study, we examine the use of rubrics to assess authentic learning exercises students complete in one-shot library sessions. Although our study analyzes outcomes from a 200-level researched writing course, our model can be adjusted for use with a variety of class sizes, lengths, and content.

We present a case study of our work at the James Branch Cabell Library of Virginia Commonwealth University (VCU) in Richmond. We used rubric-based assessment of authentic learning products from a large, sophomore-level research and writing class with dozens of sections per semester to understand what students learned in terms of information literacy and research skills. In this article, we share the assessment method we developed and how we used the process to enhance student learning and improve
our instruction program. We will also report on student achievement of learning outcomes for the librarian-led class sessions.

Literature Review

Beginning with the rise of external accreditation in the 1960s, institutions of higher education have increasingly been called upon to demonstrate their value.¹ Today, assessment has become more integral to higher education than ever before, and its importance is increasing. Academic libraries are no exception to this trend. For decades, they have responded to calls for accountability with quantitative performance data: collection size, gate counts, number and type of interactions at service desks, and so forth. Even now, the Association of Research Libraries places strong emphasis on the accumulation of resources, number of professional and support staff, and employee salaries in the process of granting membership.²

But the shape of assessment in academic libraries is changing.³ Increasingly, librarians strive to take a more active role in student learning and to be viewed as equal academic partners with university faculty.⁴ Thus, there is more interest in assessing how academic libraries impact student learning and success. Librarians increasingly look for ways to demonstrate the benefits of library services (including librarian-led instruction and other teaching initiatives) for student learning.

Assessment of Learning in Library Instruction

In their survey of assessment practices in library instruction, Karen Sobel and Cassidy Sugimoto found that librarians use a variety of approaches to assess instruction. Worksheets completed during library instruction and quizzes (both as pretests and posttests) were the most popular tools. Smaller percentages of respondents indicated that they evaluated students’ performance on specific course assignments that were completed either during or after library instruction.⁵ In a review of the literature, Andrew Walsh found multiple-choice questionnaires to be the most popular form of library instruction assessment. Analysis of bibliographies and quizzes or tests ranked next in popularity, while other forms of assessment, such as observations, self-assessments, and reviews of essays and portfolios, were used less frequently.⁶

These findings suggest a variety of approaches to student learning assessment in library instruction, but some of the most popular forms identified by Walsh and by Sobel and Sugimoto fall into the broad category of fixed-choice tests, which include quizzes, pretests and posttests, multiple-choice tests, and true-false tests.⁷ Fixed-choice tests include anything in which the student is required to select a preconstructed response, as opposed to writing his or her own answer. The literature abounds with examples of this type of assessment as measures of student learning in library instruction over

Increasingly, librarians strive to take a more active role in student learning and to be viewed as equal academic partners with university faculty.
the past couple of decades. Some of the best-known information literacy and research examples include large-scale or standardized tests, such as Project SAILS (Standardized Assessment of Information Literacy Skills), developed at Kent State University in Ohio.

Some authorities assert that fixed-choice tests, despite their prevalence, focus only on the lower levels of Benjamin Bloom’s taxonomy, a classification system used to arrange learning objectives by their levels of complexity. According to these experts, such tests measure only the lower-level objectives of knowledge, comprehension, and analysis, rather than the higher-level goals of application, synthesis, and evaluation. In other words, fixed-choice tests assess what students know instead of what they can do. These evaluations sometimes go hand-in-hand with “teaching to the test” rather than imparting to students “the essential skills of inquiry and expression.” Due to criticism of such fixed-choice tests, some recent literature calls for the use of authentic assessments, which are designed to measure real-world skills and the ability to solve complex problems. Additionally, authentic assessment looks not only at the product of learning but also at the process.

**Authentic Assessments**

One of the main benefits of authentic learning and assessments is that they are designed to have meaning beyond the specific environment in which they are completed. An authentic learning exercise for library instruction could enhance students’ research abilities in ways that extend to other classes or could even contribute to lifelong learning skills that last far beyond the students’ time in academe. Brandy Whitlock and Julie Nanavati point out that such assessments are not only useful for librarians and instructors to measure student learning but also for the students themselves: “Engaging in performative and authentic assessment encourages student investment in the learning process by making the outcomes more readily applicable and more meaningful.” For example, research logs or journals encourage students to keep track of fruitful search terms and resources, which is useful for the assignment at hand. However, the skill and the approach to searching that students take away from this type of exercise may also have bearing on future academic, professional, and personal research. Meanwhile, the artifact of the learning—the research log—provides librarians with a product that can be used to evaluate student learning.

By using authentic forms of assessment, librarians can accomplish both goals: increasing time for student-centered pedagogy and measuring student learning.

**Rubrics as a Method to Assess Learning**

Rubrics can be a useful way to assess student learning. Defined as “descriptive scoring schemes,” rubrics provide clear, concise descriptions of the performance expectations...
at each level of mastery of a product or process of learning. For the sake of brevity, we will not list exhaustively the benefits of using rubrics for assessing student learning, although we recommend exploring Megan Oakleaf’s work on rubrics in the context of libraries for additional information. That said, a few benefits of rubrics are especially worth noting for how they can be useful in course-integrated library instruction.

Creating rubrics can be a helpful way of ensuring that all teaching librarians have a shared understanding of the desired learning outcomes for a course, particularly in instances where the goals have not been explicitly stated. This is especially valuable in settings where multiple librarians share the responsibility for teaching large groups of students, as is often the case in undergraduate information literacy programs. Once the rubric is complete and artifacts of learning have been gathered, assessors can use the rubric to generate clear, descriptive data about student performance in a systematic manner, reducing subjectivity and inconsistency in assessing student work.

Finally, rubrics provide flexibility and scalability, and they can be adapted for different programs or groups of students. Rubrics are easy to modify in instances where teaching librarians determine that changes need to be made to assess a particular performance criterion. In many cases, rubrics for assessing students’ research or information literacy skills in one discipline can be used, with little or no modification for another discipline. Although evaluating student work with a rubric is more time-consuming than analyzing data from a large standardized test, issues of scale can be met by selecting a sample of many students’ work for review (a process we will detail later).

There are numerous examples of librarians using rubrics to assess students’ authentic learning products, particularly within the past decade. Authentic learning products that can be assessed with rubrics include research logs and journals, portfolios, term paper bibliographies, products from students’ participation in online tutorials, in-class worksheets, and other exercises that demonstrate students’ learning processes.

Although the use of rubrics to assess learning is well established in the library literature, we could not identify any instances in which librarians used them to assess artifacts of learning from one-shot sessions taught by multiple librarians to improve an instruction program at large. We embraced this process at VCU Libraries and will examine the benefits of this approach in terms of understanding student learning, establishing shared goals for teaching, and improving our instructional practices.

Rubrics for Assessing Learning at VCU Libraries

At VCU, we used a rubric to evaluate worksheets collected from library instruction sessions for students enrolled in a sophomore-level research and writing course called University 200: Inquiry and the Craft of Argument (UNIV 200). UNIV 200 focuses primarily on writing an argumentative research paper. The course is not typically the students’ first experience with library instruction, but it is usually the most intensive research-based
assignment they have encountered. One entire unit of the four-unit class is devoted to information literacy and research. Librarians’ involvement in UNIV 200 is part of our foundational information literacy program at VCU, and with rare exceptions, we provide at least one instruction session for each UNIV 200 section every semester. In 2013–2014, each section received at least one librarian-led session, and some got two. A team of librarians share the responsibility of teaching UNIV 200 library sessions. In 2013–2014, the year in which data for this study were collected, five librarians comprised the team.

Our assessment model for this course began with one librarian’s use of rubric-based assessment of authentic learning products to understand the extent of student learning in the library sessions she taught. In the summer semester of 2012, it became an initiative within the information literacy program at large. The data reported in this study come from the 2013–2014 academic year, following Institutional Review Board (IRB) approval to collect data for research purposes.

To assess UNIV 200 library instruction, we developed a worksheet (Appendix A) and a rubric (Appendix B) based on information literacy learning outcomes defined in concert with UNIV 200 faculty to collect evidence of student learning while providing an active educational experience for students. The information literacy learning outcomes have subsequently evolved in light of findings from this assessment project, the advancement of library search tools, curricular changes, and the new Association of College and Research Libraries Framework for Information Literacy for Higher Education. However, the learning outcomes as they were defined at the outset of fall 2013 were:

- Develop a topic-relevant search strategy related to a research question to find and use appropriate resources effectively.
- Locate source materials, including books and articles, using advanced features of Primo (VCU Libraries’ search tool) and other navigational tools to investigate and refine a research question.
- Distinguish between Primo and discipline-specific databases to select the most appropriate resources and to maximize relevant results.

Sample and Target Population

The target population was students enrolled in UNIV 200 classes in 2013–2014. We collected data from only a portion of these students because not every section of UNIV 200 was asked to complete the worksheet. Additionally, not every student who was asked to fill out a worksheet turned one in, and some students submitted a worksheet but opted out of the research study. Both of these scenarios are explained more fully later.

Librarians did not ask students to complete the worksheet in instances where the particulars of the session did not align with learning outcomes associated with the worksheet. Although UNIV 200 faculty teach their sections in accordance with a shared curriculum, there is still some variation in how individual faculty members approach the course. As a result, there is variation in UNIV 200 faculty’s expectations for what students will learn as a result of librarian-led instruction, regardless of the shared learning outcomes. Sometimes UNIV 200 sessions diverge significantly from the shared curriculum. Because our assessment model is based upon those specific learning outcomes, teaching librarians were advised to use it only when logical and appropriate.
Accordingly, not all UNIV 200 librarian-led sections were assessed. Table 1 details the number of class sessions assessed within the approved IRB date range. We did not receive IRB approval until September 23, 2013, which meant that some worksheets from classes during the 2013–2014 academic year could not be included in the study, resulting in fewer classes assessed in fall 2013 than spring 2014. All student worksheets collected in this study were completed between September 23, 2013—when we received IRB approval—and May 10, 2014, the end of the spring semester.

Sections of UNIV 200 range from 15 to 25 students each. However, some students who were asked to complete the worksheet did not do so, or finished it but did not want to part with it even for a short time. We prioritized creating a positive learning environment more than ensuring that every worksheet was turned in, so librarians avoided pressuring students to submit the worksheet if they were reluctant. Additionally, some students turned in the worksheet but opted out of the research study. We were aware that these two issues increased our risk of sampling error. We cannot predict how student worksheets that were not collected may have affected the results of our assessment. However, since our priorities in this endeavor were student learning and

---

**Table 1.**

Worksheets collected and assessed

<table>
<thead>
<tr>
<th></th>
<th>Fall 2013</th>
<th>Spring 2014</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sections offered</td>
<td>78</td>
<td>81</td>
<td>159</td>
</tr>
<tr>
<td>Sections asked to complete worksheet*</td>
<td>29</td>
<td>63</td>
<td>92</td>
</tr>
<tr>
<td>Students present in sections asked to complete worksheet*</td>
<td>628</td>
<td>1,198</td>
<td>1,826</td>
</tr>
<tr>
<td>Worksheets collected*</td>
<td>586</td>
<td>1,154</td>
<td>1,740</td>
</tr>
<tr>
<td>Worksheets sampled</td>
<td>87</td>
<td>189</td>
<td>276</td>
</tr>
<tr>
<td>Worksheets removed from sample†</td>
<td>5</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>Worksheets assessed</td>
<td>82</td>
<td>176</td>
<td>258</td>
</tr>
</tbody>
</table>

*After receiving IRB approval.
†Some worksheets were removed due to illegibility, because they were less than half completed, or because the research question was so poorly articulated that assessors could not evaluate the remainder of the worksheet.

---

We prioritized creating a positive learning environment more than ensuring that every worksheet was turned in, so librarians avoided pressuring students to submit the worksheet if they were reluctant.
the development of a positive learning environment, we considered it an appropriate sacrifice to the rigor of our methodology.

In all, we collected a total of 1,740 worksheets that were eligible to be evaluated for this study, which means that 95 percent of the 1,826 students who attended sessions after we received IRB approval turned in a worksheet. We selected a disproportionate, stratified random sample of three worksheets from each assessed class for analysis. A stratified random sample ensured that the same number of worksheets from each class was represented, enabling us to account evenly for differences in individual faculty members’ and librarians’ teaching styles, and guaranteeing that worksheets from all sections taught and assessed were included in the sample. This resulted in 276 sampled worksheets (16 percent of the worksheets collected), 18 of which were removed because they were either too illegible to evaluate or because they were less than half completed. In some cases, worksheets were removed if the research question was so poorly articulated that we could not assess the rest of the worksheet because we did not understand the student’s information need. This resulted in a final sample of 258 worksheets for analysis (15 percent of worksheets collected). A sample of this size is large enough to ensure adequate representation of the variety in student work but small enough that it is feasible for librarians to evaluate it. Table 1 details the breakdown of sections and worksheets assessed by fall and spring semesters.

Data Collection: The Worksheet

The content of the worksheet was mapped to the learning outcomes mutually agreed upon between librarians and UNIV 200 faculty. Students completing the worksheet were instructed to do the following:

- Write out their research question.
- Identify the most important words and phrases in the research question.
- Develop related search terms.
- Use Boolean operators to connect those search terms into a coherent search strategy.
- Truncate search terms when appropriate.
- Locate two scholarly sources relevant to their research questions.

Librarians had flexibility in their pedagogical approach to help students achieve the learning outcomes associated with each session. They also had flexibility in when and how they asked students to complete the worksheet. Most students filled out the worksheets during the librarian-led instruction session, but sometimes students watched a video on developing a search strategy and finished part of the worksheet before the session. In those cases, students brought the worksheets to class the day the librarian taught.

After students completed the worksheets, librarians collected and scanned them to create duplicate library copies, and then promptly returned them to the teaching professor so that students could have them back by their next class period. This was an important element of the emphasis on authentic assessment: if the artifact we used was also intended to be a learning exercise for students, it was important to return it to them quickly.
Data Analysis

We assessed student worksheets using the rubric at the end of both the fall 2013 and spring 2014 semesters. Like the worksheet, the rubric was based upon the shared learning outcomes developed by UNIV 200 faculty and librarians. It was comprised of six criteria:

- Criterion A: Student states research question or topic.
- Criterion B: Student identifies keywords/phrases from research question.
- Criterion C: Student generates similar keywords/phrases for each key concept that will enhance search strategy.
- Criterion D: Student organizes search terms and phrases in a way that reflects understanding of Boolean operators.
- Criterion E: Student truncates search terms.
- Criterion F: Student validates the discovery of scholarly resources by documenting citation information.

At the end of each semester, two reviewers completed a norming process to ensure shared understanding of the rubric. The norming process consisted of reviewing 20 to 25 sampled worksheets (the specific number varied by semester), then comparing results between the two reviewers. As necessary, the facilitator and reviewers clarified the meaning of the rubric, and in some instances, altered it for clarity. The goal of the norming process was to ensure consistent understanding of the rubric and to increase interrater reliability.

Following the norming process, the reviewers used the rubric to score the batch of worksheets at the end of fall semester (n = 87) and spring semester (n = 189). Possible scores for each criterion ranged from 0 to 2 or 0 to 3 (see the rubric in Appendix B for additional details). The maximum score a student could achieve on the worksheet was 17 points. Individual student scores were calculated by adding up the average of the two reviewers’ scores for each criterion.

After the assessment of each semester’s data, librarians used the data to spur discussion and reflection of where students seemed to be learning, and where they were not. Librarians were encouraged to adjust their teaching to improve student learning in areas where students seemed to struggle. The librarians had freedom to modify their teaching as they saw fit in terms of pedagogy, and many sought ideas and support from one another.

A final analysis in which all data for fall 2013 and spring 2014 were analyzed and compared was conducted in summer 2014 using frequency distributions, a statistical technique that displays the frequency of various outcomes in a sample. We also performed independent sample t-tests, a technique that indicates whether a difference between two groups’ averages reflects a “real” difference rather than one that occurred because of random chance in the sample selection. We established that the interrater reliability for each criterion ranged from moderate to substantial based on parameters established by
the statisticians J. Richard Landis and Gary Koch (see Table 2). Since total scores were on an interval scale as opposed to a nominal or categorical scale (which was the case for each criterion’s score), we did not use Cohen’s kappa test to assess interrater reliability of the overall scores.

Results and Discussion

Scores of the assessed worksheets (N = 258) from fall 2013 and spring 2014 were fairly normally distributed, as shown in Figure 1. The mean was 11.94, and the standard deviation—that is, how far the measurements spread out from the mean—was 2.133. The lowest total score received was 5.5, and the highest was 16.5 (with a potential maximum score of 17).

Overall, mean scores indicate that students were moderately successful in identifying the most important words in their research question and connecting their search terms with Boolean operators (see Table 3). They demonstrated some success in thinking of related search terms and finding relevant resources. The students performed least well on truncation. Although students’ average scores for Criterion A, stating their research question, were fairly good, this particular criterion was designed to indicate not the quality of the research question, but simply whether it was presented in such a way that librarians could extract enough meaning to assess the remainder of the worksheet. It was primarily used to remove worksheets from the sample when research questions were so poorly articulated that we could not judge the quality of the rest of the worksheet. Thus, the mean score for Criterion A does not reflect the quality of students’ research questions.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Description</th>
<th>Cohen's kappa*</th>
<th>p-value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Student clearly states research question.</td>
<td>0.501</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>B</td>
<td>Student identifies important words/concepts.</td>
<td>0.338</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>C</td>
<td>Student develops related search terms.</td>
<td>0.334</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>D</td>
<td>Student connects terms with Boolean operators.</td>
<td>0.689</td>
<td>0.05</td>
</tr>
<tr>
<td>E</td>
<td>Student truncates search terms as appropriate.</td>
<td>0.677</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>F</td>
<td>Student identifies two scholarly, relevant resources.</td>
<td>0.725</td>
<td>&lt; 0.05</td>
</tr>
</tbody>
</table>

* Cohen’s kappa measures the agreement between raters, taking into account the amount of agreement that would be expected by chance.
† The p-value gives the likelihood that any effect seen in the data, such as a correlation or a difference in means between groups, might have occurred by chance.
Using Rubrics to Assess Learning in Course-Integrated Library Instruction

Figure 1. Distribution of worksheet scores

Table 3.
Mean scores for each rubric criterion

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Description</th>
<th>Mean score</th>
<th>Maximum score possible</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Student clearly states research question.</td>
<td>1.82</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>Student identifies important words/concepts.</td>
<td>2.42</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>Student develops related search terms.</td>
<td>1.99</td>
<td>3</td>
</tr>
<tr>
<td>D</td>
<td>Student connects terms with Boolean operators.</td>
<td>2.45</td>
<td>3</td>
</tr>
<tr>
<td>E</td>
<td>Student truncates search terms as appropriate.</td>
<td>1.24</td>
<td>3</td>
</tr>
<tr>
<td>F</td>
<td>Student identifies two relevant scholarly resources.</td>
<td>2.02</td>
<td>3</td>
</tr>
</tbody>
</table>
We observed a statistically significant difference ($p < 0.05$) in students’ overall average scores between class sessions that were 50 minutes long ($M = 11.67$) and those that were 75 minutes long ($M = 12.33$) over the entire data collection period. The effect size, a measure of the difference between the two groups, was $d = 0.31$. While we cannot suggest a causal relationship between length of class and performance, we believe a potential reason for this to be straightforward: in 75-minute sessions, both librarians and students could devote 50 percent more time to achieving learning outcomes than in 50-minute sessions.

We noted improvement in every criterion and overall average score between fall 2013 and spring 2014, although the differences were only statistically significant for Criterion F and for overall worksheet scores (see Table 4), for which effect sizes were also calculated.

Although not statistically significant, we were pleased to see an increase in student scores on Criterion E (truncation), an area we focused on particularly to boost student learning after poor performance in fall 2013. We suspect the higher score in Criterion F (finding relevant sources) was in part due to our increased commitment to allowing students ample time to search for information during the class session, as well as significant improvement in the relevance sorting of our libraries’ discovery tool, Primo.

### Table 4.
Comparison of scores between fall and spring semester

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Description</th>
<th>Fall mean</th>
<th>Spring mean</th>
<th>p-value*</th>
<th>d-value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Student clearly states research question.</td>
<td>1.79</td>
<td>1.84</td>
<td>&gt; 0.05</td>
<td>—</td>
</tr>
<tr>
<td>B</td>
<td>Student identifies important words / concepts.</td>
<td>2.37</td>
<td>2.44</td>
<td>&gt; 0.05</td>
<td>—</td>
</tr>
<tr>
<td>C</td>
<td>Student develops related search terms.</td>
<td>1.97</td>
<td>1.99</td>
<td>&gt; 0.05</td>
<td>—</td>
</tr>
<tr>
<td>D</td>
<td>Student connects terms with Boolean operators.</td>
<td>2.44</td>
<td>2.46</td>
<td>0.05</td>
<td>—</td>
</tr>
<tr>
<td>E</td>
<td>Student truncates search terms as.</td>
<td>1.16</td>
<td>1.30</td>
<td>&gt; 0.05</td>
<td>—</td>
</tr>
<tr>
<td>F</td>
<td>Student identifies two relevant resources.</td>
<td>1.77</td>
<td>2.13</td>
<td>&lt; 0.05</td>
<td>0.31</td>
</tr>
<tr>
<td>Overall</td>
<td>Average score of all combined criteria.</td>
<td>11.67</td>
<td>12.16</td>
<td>&lt; 0.05</td>
<td>0.33</td>
</tr>
</tbody>
</table>

*The p-value gives the likelihood that any effect seen in the data, such as a correlation or a difference in means between groups, might have occurred by chance.

† The d-value is the difference in the two groups’ means divided by the average of their standard deviations.
Primo had several periods of downtime or less-than-optimal performance in fall 2013, which affected librarian-led instruction sessions.

Use of Assessment Data and Impact on Instruction

The reflective benefit of reviewing our assessment results each semester cannot be overstated. It allowed us to learn from one another and to share ideas for how to improve student learning. It also created team cohesiveness, compelling us to reconnect with learning outcomes for the sessions we offered. We shared yearly data with UNIV 200 professors and with library administration. The university used some of the assessment results in its most recent Southern Association of Colleges and Schools (SACS) reaccreditation process.

Finally, this process has given us the evidence to take action on our findings, or to “enact decisions,” an important step in Megan Oakleaf’s Information Literacy Instruction Assessment Cycle.\textsuperscript{28} The best example of how we used the data to enact decisions is our changes to how we approach truncation. We began by devoting additional attention to this concept and reconsidering how we teach it, since students found it difficult to grasp. Over time, though, after seeing only minimal improvement in students’ understanding of this concept, coupled with the more user-friendly, “Google-like” search functions of our discovery tool, we decided to eliminate truncation from UNIV 200 instruction unless the professor requested it. This decision also led us to reconsider the value of teaching Boolean operators to these relatively inexperienced researchers, and we ultimately removed that topic from our teaching repertoire as well. Instead, we focused on thoughtful keyword development and source evaluation, among other things. Additionally, based on the evidence that students found higher-quality sources when given more time to search, we streamlined the instruction portion of class to provide students with additional search time in sessions lasting only 50 minutes. In other words, our assessment results played an important role in improving and modifying our personal teaching practices as well as our instruction program.

One of the model’s greatest strengths is its ability to be applied in a variety of scenarios. It works well for one-shot instruction, for embedded instruction, or for credit-bearing classes.

Application to Other Environments

We have shared one instance of how rubric-based assessment of authentic learning products was used for UNIV 200 library sessions at VCU. However, one of the model’s
greatest strengths is its ability to be applied in a variety of scenarios. It works well for one-shot instruction, for embedded instruction, or for credit-bearing classes. It is effective for all types of content, and it fosters active and applied learning in all cases. The process of implementing this type of model consists of five steps, adapted from Oakleaf’s Information Literacy Instruction Assessment Cycle, that are applicable to a variety of circumstances:

1. Define learning outcomes.
2. Create an authentic learning exercise with an associated rubric or grading key to measure outcomes.
3. Assess learning exercises.
4. Learn from your data and make necessary changes.
5. Repeat.29

This type of assessment should be an iterative process that leads to continual improvement of library instruction programs.

Limitations

The purpose of our evaluative approach was not to “prove” that changes to librarians’ teaching styles and modifications of an instruction program led to improved student achievement of learning outcomes. Indeed, this type of assessment cannot purport any causal relationship. Using rubrics to assess authentic learning outcomes from library instruction sessions provides a real-world reflection of student learning, and that real world is often a messy one. Using our assessment process to foster discussions about how we can increase student learning in the classroom may have led to some improvement in performance that could be attributed to our greater attention to, and engagement in, improving our instructional practices. But we do not purport this to be the sole agent of change. Any number of other factors could have contributed to the change in scores in some areas, and the same could be said for the lack of improvement in scores in other areas. The upgrades to our library’s discovery tool over the course of one academic year, mentioned earlier, were just one possible contributor to students’ increased success in Criterion F.

The goal of this particular model of assessment is not to point to any one instructional practice as the reason students’ scores improved. Librarians teaching UNIV 200 classes explored a variety of methods for increasing student learning, so we cannot suggest that there was a uniform attempt to improve student learning via a particular pedagogy. Instead, our aim was to increase our own awareness of what students were learning and what they were not learning, and consider how we might change what we were doing in our program and in the classroom in light of that information.
Similarly, although students’ average total score was just under 12 out of a maximum of 17, equating to an average percentage score of approximately 70 percent, it is important to note that the purpose of this assessment was not to demonstrate student mastery, but rather to understand how we could use this evidence to improve their learning over time. We did not expect sophomore-level students to master the processes they were asked to demonstrate via the worksheet after one short librarian-led session. Our hope, instead, was to develop a snapshot of what students struggle with most so that we might consider those sticking points moving forward.

**Conclusion**

Using rubrics to assess authentic learning exercises is a useful way to evaluate student learning during librarian-led, course-integrated instruction sessions. It affords students the chance to complete an authentic learning exercise, and it gives librarians rich and useful assessment data. Additionally, the general approach can often be tweaked minimally to fit the learning outcomes of information literacy and research skills for multiple disciplines and at various levels of the curriculum.

At VCU, one of the greatest strengths of this method is that it spurred conversation among our team of teaching librarians and increased our shared understanding of what we wished to accomplish in teaching UNIV 200 students. Although this approach requires more time than asking students to complete a fixed-choice test, we believe it is worth the time because it provides students an opportunity to learn more deeply.

Laura W. Gariepy is an assistant professor at Virginia Commonwealth University in Richmond and head of teaching, learning, and information at the university’s James Branch Cabell Library; she may be reached by e-mail at: lwgariepy@vcu.edu.

Jennifer A. Stout is an instructor at Virginia Commonwealth University and a teaching and learning librarian at the James Branch Cabell Library; her e-mail address is: jastout@vcu.edu.

Megan L. Hodge is an instructor at Virginia Commonwealth University and a teaching and learning librarian at the James Branch Cabell Library; she may be reached by e-mail at: mlhodge@vcu.edu.
Appendix A

Develop a Search Strategy

Your Name: ____________________________

1. What is your research question or topic? Be as specific as you can. Circle the key concepts (the most important words or phrases).

   | Circ1: friendship       | friendship/relationships/dating |
   | AND                     | social media/social networking/facebook |
   | Circ2: social media     |                                       |

2. In the left column in the chart below, list the key concepts you circled in Number 1. In the right column, list related words or synonyms beside each key concept.

   | Circ1: friendship       | friendship/relationships/dating |
   | AND                     | social media/social networking/facebook |
   | Circ2: social media     |                                       |

   Related words/phrases in each row should be connected with AND
   Circ1: ____________________________
   Circ2: ____________________________
   Circ3: ____________________________

3. Take a look at the concepts and related phrases you listed in Number 2. In the space below, use the asterisk (*) to truncate search terms as appropriate.

4. Please provide the following citation information for at least two (2) scholarly sources you found that are relevant to your research question:

   First Source:
   Title: ____________________________
   Author: ___________________________
   Journal (if applicable): _____________
   Year of Publication: ________________

   Second Source:
   Title: ____________________________
   Author: ___________________________
   Journal (if applicable): _____________
   Year of Publication: ________________

NOTE: From Fall 2012 to Spring 2014, responses to questions on this worksheet will be used for a research study on student learning and literacy instruction. All names and identifiers will be removed from the worksheets for research purposes, and only the researchers will have access to the original worksheets. If you wish to keep your name and responses private for other reasons, please check the box at the top here.

☐ Do not wish for any responses to be included in the research study.
# Appendix B

## Rubric: Worksheet

<table>
<thead>
<tr>
<th>Criteria/Learning Outcomes</th>
<th>3 points</th>
<th>2 points</th>
<th>1 point</th>
<th>0 points</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Student states research question or topic.</td>
<td>N/A</td>
<td>Student states research question/topic clearly enough that the librarian can understand what the sought-after information should address.</td>
<td>Student states a research question/topic but does not provide sufficient detail for the librarian to understand what the sought-after information should address.</td>
<td>Student does not state a research question/topic.</td>
</tr>
<tr>
<td>B: Student identifies key words/phrases from research question.</td>
<td>Student correctly identifies all or most key concepts in research question/topic.</td>
<td>Student correctly identifies some but not all key concepts in research question/topic.</td>
<td>Student identifies concepts from research question/topic, but they are incorrect or not likely to generate relevant search results.</td>
<td>Student does not identify any key concepts.</td>
</tr>
<tr>
<td>C: Student generates similar key words/phrases for each key concept that will enhance search</td>
<td>Student produces related words/phrases for each key concept in his/her research question. The phrases represent key concepts and are different from each other.</td>
<td>Student produces related words/phrases for only a few key concepts; or, produces related words/phrases for each key concept, but the phrases are not distinct and relevant.</td>
<td>Student produces minimal related words and phrases; or, produces related words/phrases for some key concepts, but few or none are relevant.</td>
<td>Student does not produce any related words/phrases for key concepts.</td>
</tr>
<tr>
<td>D: Student organizes terms and phrases in a way that reflects an understanding of Boolean operators.</td>
<td>Organization of terms reflects limited understanding of use of the Boolean operators AND and OR.</td>
<td>Organization of terms suggests incorrect understanding of use of the Boolean operators AND and OR.</td>
<td>Student did not complete question.</td>
<td></td>
</tr>
<tr>
<td>E: Student truncates search terms.</td>
<td>Student correctly truncates search terms and phrases when appropriate.</td>
<td>Student correctly truncates some search terms/phrases when appropriate.</td>
<td>Student truncates search terms and phrases, but not correctly. Or, student correctly truncates very few words.</td>
<td>Student does not truncate search terms and phrases at all.</td>
</tr>
<tr>
<td>F: Student validates the discovery of scholarly sources by documenting citation information.</td>
<td>Student provides citation information for two scholarly sources that appear relevant to her/his question/topic.</td>
<td>Student provides citation information for only one scholarly source that appears relevant to her/his question/topic.</td>
<td>Student cites a scholarly source that very minimally pertains to his/her research question, or student cites an item that pertains to the research question but is not a scholarly source.</td>
<td>Student does not cite a source or any other item that pertains to her/his research question. Or, student does not complete question.</td>
</tr>
</tbody>
</table>
Notes


12. Ibid., 43.


20. Oakleaf, “Using Rubrics to Collect Evidence”; Oakleaf, “Dangers and Opportunities.”

21. Oakleaf, “Dangers and Opportunities.”


25. Knight, “Using Rubrics to Assess Information Literacy.”


27. Carter, “Use What You Have.”


31. Ibid., 543–45.