# Closing the Loop: Engaging in a Sustainable and Continuous Cycle of Authentic Assessment to Improve Library Instruction

Teagan Eastman
*Utah State University*, teagan.eastman@usu.edu

Kacy Lundstrom
*Utah State University*, kacy.lundstrom@usu.edu

Katie Strand
*Utah State University*, katie.strand@usu.edu

Erin Davis
*Utah State University*, erin.davis@usu.edu

Pamela N. Martin
*Utah State University*, pamela.martin@usu.edu

*See next page for additional authors*

## Let us know how access to this document benefits you.

# Closing the Loop: Engaging in a Sustainable and Continuous Cycle of Authentic Assessment to Improve Library Instruction

**Authors**

Teagan Eastman, Kacy Lundstrom, Katie Strand, Erin Davis, Pamela N. Martin, Andrea Krebs, and Anne Hedrich

# Closing the Loop: Engaging in a Sustainable and Continuous Cycle of Authentic Assessment to Improve Library Instruction

Teagan Eastman, Utah State University; Kacy Lundstrom, Utah State University; Katie Strand, Utah State University; Erin Davis, Utah State University; Pamela N. Martin, Utah State University; Andrea Krebs, University of Illinois at Urbana-Champaign; Anne Hedrich, Utah State University

## Abstract

This study demonstrates how a team of librarians sustained authentic assessment across multiple studies in order to inform changes to an information literacy curriculum. It demonstrates the cyclical and action-based nature of assessment, including closing one loop only to reopen another and begin the assessment process again, emphasizing the importance of sustainability and making changes that increase student learning. Researchers analyzed 79 English composition papers for evidence of information literacy skills, expanding upon a previous study which established information literacy skill benchmarks. Findings from the previous study led to the development of new library instruction lessons, which targeted skills students struggled with – mainly topic refinement and information synthesis. To measure the impact of the modifications, the authors used two rubrics as well as a citation analysis to identify shifts in student learning. Findings indicate that the new lessons contribute to student improvements in synthesis, topic refinement, and source variety. This study illustrates the importance of engaging in an ongoing cycle of assessment and continually making improvements to instruction practices while implementing evidence-based decisions.

*Keywords*: authentic assessment, assessment cycle, sustaining assessment, information literacy

*Eastman, et al*
*Closing the Loop*

[ ARTICLE ]

# Closing the Loop: Engaging in a Sustainable and Continuous Cycle of Authentic Assessment to Improve Library Instruction

## Introduction

Utah State University (USU) Libraries has been engaged in ongoing, cyclical assessment of library instruction in introductory writing courses. In a 2015 rubric study, USU librarians analyzed 890 papers in four key courses throughout the curriculum, including English 2010: Intermediate Writing (ENGL 2010), a second-year writing course required for all students as part of the general education curriculum (Holliday et al., 2015). The 2015 rubric study established benchmarks for student skills, which were assessed using the Association of American Colleges and Universities (AAC&U) Information Literacy VALUE Rubric (2013). It also engaged a team of librarians, library staff, and student workers at USU in learning and contributing to authentic assessment efforts.

While similar to the 2015 rubric study, the goal of the current study (which the authors will refer to as the 2018 follow-up study) was to use the 2015 benchmarks to measure the impact of the modifications made to the ENGL 2010 curriculum, thereby "closing the loop," the next step in the assessment cycle. With the 2018 follow-up study, the researchers applied not only the IL VALUE Rubric but also Lundstrom, Diekema, Leary, Harderlie, and Holliday's (2015) Synthesis Rubric and conducted a citation analysis as well.

Library instructors widely accept the value of authentic assessment; however, conducting authentic assessment in ongoing, scalable, and useful ways remains a challenge for many library instruction programs. This article describes how the researchers built on a foundation of assessment to further understand the needs and abilities of students, and to make programmatic changes and decisions informed by a continuous cycle of assessment. The methods used are scalable and relevant to librarians interested in understanding their impact on student learning and continuously seeking to improve it.

## Background

USU Libraries rely on a course- and curriculum-based library instruction program. A core piece of the program is integration with ENGL 2010. Each semester a librarian is partnered with every section offered (approximately 70). ENGL 2010 instructors bring their classes to

[ ARTICLE ]

the library for two or three sessions focused on supporting the major assignment, a persuasive research essay.

Past collaborations with ENGL 2010 instructors varied widely from instructor to instructor and from librarian to librarian. It became clear that some instructors and librarians would benefit from more structured activities tied to learning outcomes when planning their library sessions. The 2015 rubric study concluded that "students struggled most in categories that required critical thinking, including evaluating information, synthesizing information, and using information effectively in their writing" (Holliday et al., 2015, p. 178). In response, librarians created two new lessons that targeted students' need to develop the ability to refine a topic and synthesize information, titled "Narrowing a Topic" and "Synthesis," respectively (for lesson plans, see: https://libguides.usu.edu/2010lessons/2015lessons). While some ENGL 2010 sections adopted one lesson or the other, ten sections used both of the new lessons, albeit with some adaptations.

## Literature Review

Assessment is essential as academic libraries consistently work to determine their impact and demonstrate their value (Mery, Newby, & Peng, 2012). Current assessment practices employed by academic librarians take on multiple formats such as citation analysis, rubrics, student portfolios, surveys (Larsen, Izenstark, & Burkhardt, 2010, p. 64). Librarians have diversified their assessment practices as university stakeholders are no longer impressed with circulation statistics and door counts; instead, they want more authentic assessment tied to actions that improve services. They want to know what skills students develop as a result of interactions with the library and how students are retaining and applying those skills (Oakleaf, 2008). Additionally, the assessment interests of academic libraries have shifted since the adoption of the *Framework for Information Literacy for Higher Education* (Association of College and Research Libraries [ACRL], 2015) and initiatives such as ACRL's Assessment in Action (AiA) program. Assessment paves the way to improved pedagogy, celebrating learning achievements as well as diagnosing existing problem areas (Oakleaf, 2008). However, many libraries are still learning how to practice authentic assessment in sustainable, practical ways; this study is an examination of how one library is approaching this challenge.

In a review of the information literacy assessment literature, Erlinger (2018) reported that 41% of the studies included some form of authentic assessment. Authentic assessments are "those that challenge students to apply what they have learned in real-life situations…. [and]

it is generally agreed that [they] are most effective when tied to an existing, graded course assignment" (Erlinger, 2018 p. 452). In the literature, authentic assessment has been used to evaluate information literacy across campus (Gola, Ke, Creelman, & Vaillancourt, 2014), assess online library instruction (Alverson, Schwartz, & Shultz, 2018), and gauge learning in library one-shots (Gariepy, Stout, & Hodge, 2016). While authentic assessment can be worthwhile and eye-opening, Erlinger admitted that it is time-consuming and requires a high-degree of collaboration with instruction faculty (2018, p. 452).

Librarians have often employed rubrics as an authentic assessment tool to score student work, noting the limitations of other methods such as citation analysis (Rosenblatt, 2010). The benefits of using a rubric are two-fold. First, they provide students with a road map of what they need to incorporate in an assignment, give instant feedback, and provide context for grading (Belanger, Zou, Mills, Holmes, & Oakleaf, 2015). Second, for librarians, constructing a rubric can push an instruction department to articulate their instruction goals in a way that aligns with an institution's mission (Rosenblatt, 2010). Rubrics also "communicate agreed-upon learning values, focus on standards and concepts, align with educational theory, and provide results that can be applied to improve instruction" (Belanger et al., 2015, pp. 623–624).

Previous research and projects were helpful in developing and implementing scoring rubrics for this study. Gola, Ke, Creelman, and Vaillancourt (2014) described a campus-wide effort to develop an information literacy rubric and noted the need for "more case studies and best practices" in the library literature (p. 133). Erlinger's literature review covered many rubric assessments, noting that "rubrics must be tailored to the situation" to be effective (2018, p. 453). Project RAILS, an Institute of Museum and Library Services (IMLS) grant-funded project, facilitated a large-scale rubric assessment involving nine institutions, which led to a set of recommendations or best practices from several participants. This 2018 follow-up study built on the four recommendations, including "(1) building successful collaborative relationships, (2) developing assignments, (3) creating and using rubrics, and (4) using assessment results to improve instruction and assessment practices" (Belanger et al., 2015, p. 624).

In the original 2015 rubric study, the researchers used the AAC&U Information Literacy VALUE Rubric, which is based on the ACRL *Information Literacy Competency Standards for Higher Education* (2000). The ACRL Board's decision to rescind the *Standards* in June 2016 has had consequences for library instruction programs that relied on the *Standards* or that

[ ARTICLE ]

were in the process of shifting their assessment processes to include the threshold concept-based *Framework for Information Literacy for Higher Education* (ACRL, 2015). Oakleaf, a strong proponent of authentic assessment, noted that the *Framework* and authentic assessment can go hand-in hand, asserting that "threshold concepts are very well suited to learning outcomes assessment, as long as the assessments permit the use of authentic assessment approaches" (Oakleaf, 2014, p. 511). However, many institutions had used the *Standards* to assess their programs and student work on a granular level. Although the *Framework* has been used to develop new programmatic assessment plans (Gammons & Inge, 2017), Witek (2015) acknowledged the dilemma for institutions whose long-term assessment plans drew from the original *Standards*. However, Witek and others have emphasized the connections between the *Standards* and the *Framework* (Krug, 2016). In developing methods for the 2018 follow-up study, the researchers felt the comparison opportunities presented by the 2015 rubric study, the connections between the *Standards* and the *Framework*, and the planned inclusion of additional rubrics still warranted using the IL VALUE Rubric. Synthesis is addressed in both documents, with the *Framework* categorizing it as "ideas gathered from multiple sources" as a knowledge practice for the "Research as Inquiry" frame, whereas the *Standards* states it as "[using] information effectively to accomplish a purpose" (Standard 4). Both rubrics attempt to clarify how effective information synthesis can be identified.

In addition to using the revised IL VALUE Rubric from the 2015 rubric study (Holliday et al., 2015), the researchers used a synthesis rubric adapted and developed by a team of USU researchers. After creating an intervention to help students improve their ability to synthesize information, Lundstrom et al. (2015) created the Synthesis Rubric to measure students' varying levels of information integration in their final essays. In this 2018 follow-up study, the researchers used the Synthesis Rubric to score papers for evidence of synthesis, which was a focus of IL sessions for the sample courses. Although Lundstrom et al. used a different intervention to teach the synthesis process, the researchers for the 2018 follow-up study felt the Synthesis Rubric would also be useful in assessing information synthesis in the papers of students who had received the revised library instruction sessions.

This 2018 follow-up study is inspired by the cyclical nature of assessment. In the early 2000s, Maki described the need to create cyclical assessment plans to create "an organic process of discovering how and what and which students learn" (2002, p.13). Oakleaf outlined a library-specific Information Literacy Instruction Assessment Cycle (ILIAC) to help guide assessment work (Oakleaf, 2009). Additionally, action research, recently championed by ACRL's Value of Academic Libraries AiA initiative, has been described in

the literature as cyclical assessment that "puts action at its core and seeks to create change and study the results" (Vezzosi, 2006, p. 290). In the tradition of action research, this study describes one step in the never-ending assessment cycle – to investigate whether the "closed loop," or modifications in response to previous assessments, impacted student learning.

## Methods

### Sample

Scalability was a major factor in deciding the sample size for the research study. In total, 79 final persuasive research papers were submitted. Three papers were discarded because they did not include a works cited list, leaving a total recorded sample of 76 student papers. Although the sample is significantly less than the 890 papers scored in the 2015 rubric study, the narrower focus of this research study (determining if the revised lessons were improving the synthesis skills for English 2010 students versus setting benchmarks of IL skills across the entire curriculum) warranted decreasing the sample size. Additionally, the smaller sample size allowed for a more reasonable workload for librarians assisting with the research.

Although the length and stipulations for the final assignment varied slightly from section to section, all students were required to write a six- to ten-page paper with a research component and persuasive angle. The researchers targeted ten traditional (taught face-to-face) sections of ENGL 2010 that used both lessons designed to address troublesome concepts for students that emerged in the first study, narrowing a topic and synthesis. While most semesters have a 100% participation rate from ENGL 2010 courses in the library instruction program, not all of the courses used the two new lessons. Some courses used only one of the established lessons, some significantly adapted the lessons, and some instructors taught the lessons themselves, using the library sessions to provide students time to conduct research for their assignment with their librarian and instructor present. The researchers pulled a sample of five to seven papers from each participating section, based on how many students agreed to participate and followed up by emailing a copy of their final paper.

The researchers stripped the 79 submitted papers of identifying information and sent the student identification numbers to the Office of Analysis, Assessment and Accreditation to collect demographic data. Five of the identification numbers did not match up in the university student records system, leaving 74 students for the demographic analysis. The

sample population was comparable to the overall ENGL 2010 population in terms of gender breakdown, with a slightly higher percentage of males in both. However, the study sample population had a significantly higher percentage of freshmen and sophomores, with approximately 24% fewer seniors (see Appendix A, Table A1). Data comparing academic achievement suggest a slight difference, with students in the study sample having a higher average admission index and a slightly higher average cumulative GPA than the ENGL 2010 population (see Appendix A, Table A2).

## Rubric Rating Procedures

A team of six librarians scored students' final persuasive research papers. After reading six papers together and having all scorers use both the IL VALUE and Synthesis rubrics, the researchers decided to split the team into two groups of three. One group was assigned to score all the papers with the Synthesis Rubric and one group was assigned to score using the IL VALUE Rubric. The three team members in the IL VALUE Rubric group participated in the 2015 rubric study and therefore had previous experience scoring papers with this rubric. The other group had not scored using the Synthesis Rubric prior to this study.

The researchers used Krippendorf's alpha (KA) to determine inter-rater reliability so that they could eventually split up the 76 papers among each rubric team member. As in the 2015 rubric study, they chose this test because it applies to ordinal data and is effective across multiple coders. The researchers used the web-based calculator ReCal for Ordinal, Interval, and Ration Data (OIR) to determine the KA. The goal was to reach a KA level of 0.61, indicating good agreement.

## IL VALUE Rubric Team Process

The IL VALUE Rubric team engaged in the process of scoring six papers at a time and checking inter-rater reliability. Although the inter-rater scores remained relatively consistent and improved in each category, they did not meet the goal of good agreement (KA .61 or higher) by the fourth round (see Appendix B, Table B1). Throughout the rounds, the team discussed and reached consensus on scores that had not matched across two scorers. After the fourth round, by which time nearly half of the sample had been read and scored, the team decided that the most consistent and rigorous scoring process would be to continue with all three team members scoring all 76 papers. When at least two of the scorers for each category agreed, the mode score was used. When no scorers agreed, the median score was recorded (see Appendix B, Table 2).

## Synthesis Rubric Team Process

Because the Synthesis Rubric team had not scored papers using this rubric previously, they started by scoring five non-sample papers and discussing each score to align their reasoning. They achieved a good inter-rater agreement score on categories C and E in this round of scoring. Next, the Synthesis Rubric team read and scored five papers from the study sample. The combined inter-rater reliability of the first five papers and the sample papers achieved a score of .61 in all categories except for category A, which rated .482 (see Appendix B, Table 3).

The team determined that these scores, while achieving only fair agreement in category A, were good enough to move forward in splitting the papers and scoring individually among their team of three. This decision allowed each person to score approximately 25 additional papers, instead of all 76. For the five study sample papers scored during round two, the team used the score that two of the three members agreed upon.

## Citation Analysis Procedures

For the citation analysis part of the study, a team of five library staff and student workers analyzed the bibliographies of the sample papers, evaluating the bibliographies as a whole, the individual citations, and the journals cited. The team recorded the number of citations listed in the bibliographies and gave each bibliography a score between 1 and 5 for source variety using a rubric (see Table 1). For example, if a bibliography cited three scholarly journal articles and two books it would receive a source variety score of 2.

**Table 1: Citation Analysis Source Variety Rubric**

| Score | | | | |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| No variety of sources; cites only one type of source | Poor variety of sources; cites two types of sources | Adequate variety of sources; cites three types of sources | Good variety of sources; cites four types of sources | Excellent variety of sources; cites more than four types of sources |

The team collected information about each source cited, including the publication year, author, title, resource type, resource name, database, and the student paper number. For scholarly journal articles, the team further recorded the journal title, peer review status, and up to three subject classifications using *Ulrich's Periodicals Directory.*

## Findings

IL VALUE Rubric Comparative Findings

*Category A: Defines the Extent of Information Needed*

Category A examined how well students define the information needed. The scorers carefully considered the scope of each student's research questions and key concepts. To score highly in this category, papers needed to have a well-articulated thesis statement or research question with a manageable scope – neither too broad nor too narrow. The 2018 study scores were higher than the scores from the 2015 rubric study, in which most students scored in the mid-range of 1.5 and 2. In 2015, 71.9% scored 2 or higher compared to 84.2% scoring 2 or higher in 2018. This increase included a shift in the 2018 scores towards the 3 – 3.5 score range, a 22% increase from 2015 (see Table 2). Paper topics that warranted a score of 3 selected a topic that supported the scope of the assignment, including addressing the most relevant key concepts. One student who scored a 3 focused on the importance of book design to its marketing and message; another explored the legislation of psychedelic drug research. Both authors demonstrated the ability to narrow a topic effectively, a skill emphasized in the first library session lesson.

**Table 2: IL VALUE Rubric Findings**

| Category A: Defines the Extent of Information Needed | | | | | |
|---|---|---|---|---|---|
|  | 0 – 0.5 | 1 – 1.5 | 2 – 2.5 | 3 – 3.5 | 4 |
| 2015 | .2% | 28.1% | 60.5% | 10.8% | .6% |
| 2018 | 0.0% | 15.8% | 51.3% | 32.9% | 0.0% |

*Category B: Access the Needed Information*

Category B evaluated whether students accessed the needed information. The scorers looked at quality and appropriateness of sources as well as variety and relevance. In this category, students' work again indicated improvement over the 2015 scores, with the 3-3.5 range jumping from under 8% to over 30% (see Table 3). Papers that scored poorly often lacked relevant sources and source variety. For example, one student wrote about the benefits of choosing the medical profession as a career path. The student scored a 1 because the sources used were only somewhat relevant, didn't include a solid background source, and relied entirely on websites. Another student, who scored a 3, wrote about prison systems and

included a range of sources including scholarly articles, magazine articles, and a few background sources.

**Table 3: IL VALUE Rubric Findings**

| Category B: Access the Needed Information | | | | | |
|---|---|---|---|---|---|
| | 0 – 0.5 | 1 – 1.5 | 2 – 2.5 | 3 – 3.5 | 4 |
| 2015 | .6% | 54.1% | 36.8% | 7.6% | 0.6% |
| 2018 | 0.0% | 22.3% | 46.1% | 31.6% | 0.0% |

*Category C: Evaluate Information and its Sources Critically and Thoroughly*

Category C investigated how thoughtfully students considered their sources. Students scored higher if they questioned assumptions and presented context for their sources and arguments. While the percentage of students with the highest score awarded (3-3.5) decreased from 5.7% in 2015 to 2.6% in 2018, many more students achieved a score of 2 or above in the 2018 scores, and fewer received the lowest score range possible (see Table 4). According to the IL VALUE Rubric, students who receive a 2 are able to "question some assumptions and identify several relevant contexts when presenting a position…[but] do not yet understand the value of different kinds of evidence or ways of knowing by discipline." The researchers felt that the assignment itself did not directly require this type of elaboration on their use of evidence and acknowledgement of opposing arguments, thus fewer than half were able to demonstrate this level of critical thought and evaluative ability.

**Table 4: IL VALUE Rubric Findings**

| Category C: Evaluate Information Critically and Thoroughly | | | | | |
|---|---|---|---|---|---|
| | 0 – 0.5 | 1 – 1.5 | 2 – 2.5 | 3 – 3.5 | 4 |
| 2015 | 15.0% | 52.6% | 26.6% | 5.7% | 0.0% |
| 2018 | 1.3% | 56.6% | 39.5% | 2.6% | 0.0% |

*Category D: Use Information Effectively to Accomplish a Specific Purpose*

Category D explored how effectively students organized their sources, including the ability to synthesize information. As in the Synthesis Rubric, the scorers evaluated whether students mixed and mingled their sources, grouping information idea by idea rather than reporting out each source separately. This category reflects the most significant improvement with about 50% achieving a score of 3 or better, compared to a little over 10%

[ ARTICLE ]

*Eastman, et al*
*Closing the Loop*

in the previous study (see Table 5). The major shift in student behaviors included organizing sources by idea, putting them in conversation with one another. More often, a paragraph had more than one source instead of dividing sources into separate paragraphs, which was a major focus of one of the revised lessons.

**Table 5: IL VALUE Rubric Findings**

| Category D: Use Information Effectively to Accomplish a Purpose | | | | | |
|---|---|---|---|---|---|
| | 0 – 0.5 | 1 – 1.5 | 2 – 2.5 | 3 – 3.5 | 4 |
| 2015 | 3.2% | 39.1% | 47.1% | 10.5% | 0.4% |
| 2018 | 1.3% | 10.5% | 38.2% | 47.4% | 2.6% |

*Category E: Access and Use Information Ethically and Legally*

Category E evaluated how well students cited outside sources, and the scores represent a marked decline in appropriate citations compared to the previous study (see Table 6). There were more instances of students presenting evidence as known facts without providing citations. Often, sources were missing from the reference page, or there were mistakes in attribution. The scorers did not evaluate the accuracy of the citation style but instead focused on the presence and consistency of attribution.

**Table 6: IL VALUE Rubric Findings**

| Category E: Access and Use Information Ethically and Legally | | | | | |
|---|---|---|---|---|---|
| | 0 – 0.5 | 1 – 1.5 | 2 – 2.5 | 3 – 3.5 | 4 |
| 2015 | 0.4% | 3.0% | 8.9% | 26.8% | 60.5% |
| 2018 | 0.0% | 19.7% | 64.5% | 14.5% | 1.3% |

Synthesis Rubric Findings

*Category A: Source Variety*

This category involved analyzing students' ability to incorporate research from a variety of sources. Students that struggled in this category tended to include sources that covered only one point of view. Sixty percent of students used sources that covered at least two or more perspectives. Only 15.8% of students were ranked as "Advanced" in this category because the majority of students failed to include sources to support opposing viewpoints (see Table 7). A few students did mention opposing viewpoints, but they did not back up their statements

*Eastman, et al*
*Closing the Loop*

[ ARTICLE ]

with evidence. Performance in this category suggests that students still struggle to identify a diverse range of information.

*Category B: Using Information from Sources Correctly*

The majority of ENGL 2010 students struggled to integrate outside sources correctly within their papers. Most students relied heavily on direct quoting. Students ranked as "Needs Improvement" only incorporated sources through direct quoting, which were often irrelevant, serving no clear purpose to support their thesis. In other cases, students used direct quotations to back up statements but provided no commentary or synthesis. The majority, or 63% of students, fell into the "Developing" category as they began to paraphrase and summarize ideas. These students failed to consistently make explicit connections between the ideas and their thesis. Twenty-four percent of students demonstrated success in using information effectively (see Table 7). These students rarely utilized direct quotations and made clear and explicit connections between sources and their thesis. For example, one student paper on the importance of raising mental health awareness for LGBTQ youth, synthesized information from multiple resources and explicitly connected the information back to the thesis statement. Overall, results from this category suggest a need for targeted instruction on how to successfully integrate and relate sources back to main ideas.

*Category C: Identifies Conversations Among Information from Different Sources*

This category revealed that students struggle to identify scholarly conversations. Students failed to consistently indicate relationships among sources, and it was difficult to determine how they supported the thesis. The 28.9% of students ranked as either "Unacceptable" or "Needs Improvement" mainly cited one source per paragraph and failed to group sources by idea (see Table 7). Students in the "Developing" category began to make explicit connections between sources but often left the reader to infer patterns and relationships. Only students in the "Advanced" category mentioned contradictory viewpoints.

*Category D: Organizes Sources in a Meaningful Way*

Students who received low scores in this category tended to organize their papers by source rather than topic or idea. These students often included information irrelevant to their thesis, and they failed to organize their information to create an impactful argument. Students who demonstrated moderate success in this category began to organize their sources and arguments in a way that revealed some patterns, but they failed to be consistent. Additionally, these students did not provide adequate analysis for their readers. High-

[ ARTICLE ]

performing students organized their sources logically to make clear connections (see Table 7).

*Category E: Analyzes Sources to Create Something New or Draw Conclusions and Make Generalizations*

This category received both the highest percentage of "Advanced" (27.6%) and "Needs Improvement" (28.9%) scores (see Table 7). Students who received high scores in this category had well-reasoned conclusions that were supported by multiple sources and critical analysis. Students who received low scores failed to provide critical analysis of their sources, resulting in unclear conclusions. These students often included details irrelevant to their thesis which distracted from their discussion. For example, in a paper arguing for the benefits of outdoor recreation in maintaining good mental health, the student included sources arguing the benefits of animal companionship, which distracted from their argument. The students who received "Developing" scores failed to support their conclusions with multiple perspectives.

**Table 7: Synthesis Rubric Scores per Category**

| Category | Unacceptable | Needs Improvement | Developing | Advanced/ Mastery |
|---|---|---|---|---|
| A: Source Variety | 0.0% | 23.7% | 60.5% | 15.8% |
| B: Uses Info Effectively | 0.0% | 13.2% | 63.2% | 23.7% |
| C: Identifies Conversations | 1.3% | 27.6% | 53.9% | 19.7% |
| D: Organizes Sources | 0.0% | 26.3% | 53.9% | 19.7% |
| E: Analyzes Sources | 0.0% | 28.9% | 43.4% | 27.6% |

## Citation Analysis Findings

The citation analysis provided an additional way of evaluating students' bibliographies. Each bibliography was scored on a scale of 1 to 5, which directly reflected the number of source types used in each paper. Source types included scholarly journals, web pages, news sources, books, magazines, etc. The average variety score was 3.31, meaning that the majority of papers cited at least three different source types. Twenty-five student papers had a variety score of 4, and 14 papers had a variety score of 5, while only nine had a variety score of 1 (see Table 8).

*Eastman, et al*
*Closing the Loop*

[ ARTICLE ]

**Table 8: Source Variety Citation Analysis Scores**

| Score | Number of Papers |
|-------|------------------|
| 1 | 9 |
| 2 | 12 |
| 3 | 17 |
| 4 | 25 |
| 5 | 14 |

It should be clarified that a low variety score does not necessarily mean the student conducted inadequate research. Over 80% of the papers that received a variety score of 1, cited primarily scholarly sources. In comparison, papers with a variety score of 5 had a much higher usage of web pages and less than 30% of the sources cited were scholarly journals. This finding demonstrates that source variety is prevalent in ENGL 2010 papers; however, the analyses using the IL VALUE Rubric and Synthesis Rubric provide more insight into the actual quality and relevance of the sources.

**Table 9: Types of Resources Cited**

| Resource Type | Percentage of Citations |
|---------------|-------------------------|
| Scholarly Journal | 41.7 |
| Web Page | 21.4 |
| Magazine | 9.8 |
| News Source | 8.0 |
| Book | 7.8 |
| Video/Film | 3.3 |
| Other | 3.3 |
| Primary Source | 1.5 |
| Personal Communication | 1.3 |
| Government Document | 0.7 |
| Reference Resource | 0.6 |
| Wikipedia | 0.4 |

Table 9 shows the types of resources used, with 41.7% of sources cited being from scholarly journals, followed by web pages at 21.4%. These data suggest that students are extensively engaging with scholarly materials; however, they are exploring other types of sources as

[ ARTICLE ]

*Eastman, et al*
*Closing the Loop*

well. For example, magazines (9.8%), news sources (8.0%), and books (7.8%) represent a considerable amount of the sources cited.

## Discussion

### Demonstrated Improvements in Student Work

Authentic assessments provided the researchers with a clear understanding of students' information literacy skills and strong evidence for where to focus future improvements to library instruction rather than making adjustments just based on what librarians thought was occurring. Despite the benefits, authentic assessments are time-consuming and many libraries struggle to sustain their projects. By building on previous assessment results, the researchers gained a more sustainable method for continually gauging student learning on a smaller scale. The previous findings also allowed for comparative analysis, which enabled the researchers to close the loop by measuring the impact of the instruction interventions introduced to the ENGL 2010 curriculum in 2016. It confirmed assumptions that students' abilities to refine their topics and synthesize information had significantly improved. The additional data gathered through the use of the Synthesis Rubric and citation analysis can be used as future benchmark measurements as librarians target those specific areas to refocus instruction efforts. This process of establishing benchmarks and conducting comparative analyses can be adopted by any institution hoping to build a culture of sustainable assessment.

While the results of the two rubrics and citation analyses viewed separately reveal important information about student behaviors and skills, cross-analysis of the data illuminates important connections. The findings indicate that while students on average use three types of sources, most students do not yet understand the value of different types of evidence nor do they evaluate their sources for credibility and reliability. Additionally, the majority of students attempt to cover different perspectives but are not yet including opposing viewpoints. This finding points to a need for focused instruction on source evaluation and incorporating opposing views.

### Continued Areas of Challenge for Students

Students are attempting to use synthesis in their papers as evidenced by the average score of 2.5 in category D (using information effectively to accomplish a purpose) of the IL VALUE Rubric. However, they have not yet mastered the skills necessary to successfully execute synthesis in order to enhance their arguments. Using the Synthesis Rubric, the researchers isolated the specific aspects of synthesis that challenged students. The data show the

majority of students are able to place sources in conversation with one another in meaningful and purposeful ways. However, they are unable to effectively draw conclusions by adding their own analysis. This finding indicates that the revised lessons are helping students to improve some aspects of information synthesis, but the higher-level work of using their own voice to create meaning remains a challenge. The use of multiple rubrics on a single sample is a scalable method for gaining a richer understanding of student abilities.

## Benefits and Uses of Assessment for Librarians

This study demonstrates a method for using assessment to build successful collaborative relationships. The process of coordinating ten team members, which included librarians, library staff, and student workers, to conduct a collaborative assessment using three rubrics was challenging but worthwhile. This process served not only as a mechanism for understanding student learning but also as an opportunity for community building. Each small group needed to work together to reach a mutual understanding of their assigned rubrics, analyze papers, and make meaning out of the results. Additional challenges presented themselves as teams struggled to reach consensus and good inter-rater reliability, forcing one team to make the decision to read all papers in the sample. After the individual groups completed their sections, the entire team came together to discuss results, identify trends, and determine implications for the ENGL 2010 program. This deep immersion into student work as an instructional team allowed for multiple perspectives and interpretations of student learning to be shared and was an opportunity for reflecting upon teaching assumptions.

Admittedly, research projects like this one do take time and resources, but there are ways to adapt and scale back projects like these, and to think carefully about use and design of the assessment so that the potential gains are long-term:

- Start with a sample of student work from one class. Keep it small at the beginning.
- Select a class that has the potential for change. Collaborate with instructors who are open to using what you learn to improve and change future instruction.
- Start with consensus. Score as a team (or a pair) and discuss your scoring choices.
- Adapt the rubric as needed; it is not set in stone and can often be improved. Keep careful track of what language you can add or revise to help the scorers better identify each level of the rubric.
- Acknowledge the gains. While parts of the process may be slow moving, focus on what you are learning and how you will use it.

[ ARTICLE ]

- Consider how your project will fit into future assessments. How does this assessment complement other methods you are using?

Although the Writing Program leadership is in flux and a new curriculum is being implemented, the researchers shared this assessment data in order to advocate for the value of continued inclusion of a dedicated synthesis lesson as part of the ENGL 2010 IL curriculum. The prevalence of source variety in students' papers demonstrates that students are more interested in the information itself rather than the defining source type. Synthesis can more effectively occur when students and instructors keep an open mind to the types of sources cited in the paper rather than forcing scholarly sources onto a topic that does not lend itself to scholarly research or analysis. The assessment data from the 2018 follow-up study help support the librarians' claims and engage faculty in a discussion about students' synthesis skills.

In order to use the assessment data to improve instruction practices further, the researchers are currently collaborating with the Writing Program to introduce new modifications. As the findings suggest, students are struggling with evaluating sources and recognizing different contexts. Thus, the researchers are in the development phase of creating an evaluation-focused session that can be taught in conjunction with the narrowing a topic and synthesis lessons. Additionally, the assessment results are being used to modify the narrowing a topic and synthesis lessons to better address skills such as incorporating multiple perspectives and using one's own perspective to support argument analysis.

## Conclusion

The findings indicate that the two lessons created as a result of the 2015 rubric study are contributing to better synthesis and improvements in other information literacy skills, such as refining scope and selecting a variety of sources. However, as the results from the 2018 follow-up study reveal, a single round of assessment and modifications cannot fill every gap in student information literacy skills. This study demonstrates the importance of engaging in a continuous cycle of assessment combined with continual conversations with faculty about their perceptions of what their students are learning.

This study has a few limitations worth noting. The student sample showed differences in achievement and class level when compared to the general ENGL 2010 population. Additionally, five student identification numbers could not be traced to enrolled students. The study also uses the IL VALUE Rubric, which is based on the rescinded IL *Standards.*

Despite these limitations, this research still has value in demonstrating shifts in student learning based on using authentic assessment methods.

These assessments instill greater confidence in the decision to shift the focus of library instruction sessions based on the findings and can be used to demonstrate added value to stakeholders. Combined with other targeted assessments in USU Libraries Instruction Assessment Plan, this research helps inform decisions about what students need and how changes to practices can better meet those needs. In looking forward, the researchers will continue to focus on assessments that are useful and actionable in understanding how to use limited resources and time in ways that will contribute the most to student learning.

## References

Association of American Colleges and Universities. (2013). *Information literacy VALUE rubric.* Retrieved from http://www.aacu.org/value/rubrics/information-literacy

Association of College and Research Libraries. (2000). *Information literacy competency standards for higher education.* Retrieved from https://alair.ala.org/handle/11213/7668

Association of College and Research Libraries. (2015). *Framework for information literacy for higher education.* Retrieved from http://www.ala.org/acrl/standards/ilframework

Alverson, J., Schwartz, J., & Shultz, S. (2018). Authentic assessment of student learning in an online class: Implications for embedded practice. *College & Research Libraries.* Preprint. Retrieved from https://crl.acrl.org/index.php/crl/article/view/16889/18535

Belanger, J., Zou, N., Mills, J. R., Holmes, C., & Oakleaf, M. (2015). Project RAILS: Lessons learned about rubric assessment of information literacy skills. *portal: Libraries and the Academy, 15*(4), 623-644. https://doi.org/10.1353/pla.2015.0050

Erlinger, A. (2018). Outcomes assessment in undergraduate information literacy instruction: A systematic review. *College & Research Libraries, 79*(4), 442-479. https://doi.org/10.5860/crl.79.4.442

Gammons, R. W., & Inge, L. T. (2017). Using the ACRL *Framework* to develop a student-centered model for program-level assessment. *Communications in Information Literacy, 11*(1), 168-184. https://doi.org/10.15760/comminfolit.2017.11.1.40

[ ARTICLE ]

Gariepy, L. W., Stout, J. A., & Hodge, M. L. (2016). Using rubrics to assess learning in course-integrated library instruction. *portal: Libraries and the Academy, 16*(3), 491-509. https://doi.org/10.1353/pla.2016.0043

Gola, C. H., Ke, I., Creelman, K. M., & Vaillancourt, S. P. (2014). Developing an information literacy assessment rubric: A case study of collaboration, process, and outcomes. *Communications in Information Literacy, 8*(1), 131-144. https://doi.org/10.15760/comminfolit.2014.8.1.157

Holliday, W., Dance, B., Davis, E., Fagerheim, B., Hedrich, A., Lundstrom, K., & Martin, P. (2015). An information literacy snapshot: authentic assessment across the curriculum. *College and Research Libraries, 76*(2), 170-187. https://doi.org/10.5860/crl.76.2.170

Krug, E. (2016). Map of ACRL *Standards* to ACRL *Framework* and King core competencies. Retrieved from https://prezi.com/nqwmqvlrl_ks/map-of-acrl-standards-to-acrl-framework-and-king-core-competencies/

Larsen, P., Izenstark, A., & Burkhardt, J. (2010). Aiming for assessment: Notes from the start of an information literacy course assessment. *Communications in Information Literacy, 4*(1), 61-70. https://doi.org/10.15760/comminfolit.2010.4.1.88

Lundstrom, K., Diekema, A. R., Leary, H., Haderlie, S., & Holliday, W. (2015). Teaching and learning information synthesis: An intervention and rubric based assessment. *Communications in Information Literacy, 9*(1), 60-82. https://doi.org/10.15760/comminfolit.2015.9.1.176

Maki, P. L. (2002). Developing an assessment plan to learn about student learning. *Journal of Academic Librarianship, 28*(1/2), 8-13. http://doi.org/10.1016/S0099-1333(01)00295-6

Mery, Y., Newby, J., & Peng, K. (2012). Performance-based assessment in an online course: Comparing different types of information literacy instruction. *portal: Libraries and the Academy, 12*(3), 283-298. https://doi.org/10.1353/pla.2012.0029

Oakleaf, M. (2008). Dangers and opportunities: A conceptual map of information literacy assessment approaches. *portal: Libraries and the Academy, 8*(3), 233-253. https://doi.org/10.1353/pla.0.0011

Oakleaf, M. (2009). The information literacy instruction assessment cycle: A guide for increasing student learning and improving librarian instructional skills. *Journal of Documentation, 65*(4), 539-560. https://doi.org/10.1108/00220410910970249

Oakleaf, M. (2014). A roadmap for assessing student learning using the new *Framework for information literacy for higher education. Journal of Academic Librarianship, 40*(5), 510-514. https://doi.org/10.1016/j.acalib.2014.08.001

Rosenblatt, S. (2010). They can find it, but they don't know what to do with it: Describing the use of scholarly literature by undergraduate students. *Journal of Information Literacy, 4*(2), 50-61. https://doi.org/10.11645/4.2.1486

Vezzosi, M. (2006). Information literacy and action research: An overview and some reflections. *New Library World, 107*(7/8), 286-301. https://doi.org/10.1108/03074800610677272

Witek, D. (2015, January 27). "Sunrise, sunset": A reflection on assessment and the *Framework for information literacy for higher education* [Blog post]. Retrieved from https://acrlog.org/2015/01/27/sunrise-sunset-a-reflection-on-assessment-and-the-framework-for-information-literacy-for-higher-education/

## Appendix A: Study Sample and ENGL 2010 Population Comparisons

**Table A1: Gender and Class Level**

|  | Demographic | Headcount | % of Total | Headcount | % of Total |
|---|---|---|---|---|---|
| Gender | Female | 36 | 48.65% | 7,279 | 47.64% |
|  | Male | 38 | 51.35% | 8,001 | 52.36% |
| Class Level | Freshman | 32 | 43.24% | 4,448 | 29.11% |
|  | Sophomore | 29 | 39.19% | 3,423 | 22.40% |
|  | Junior | 10 | 13.51% | 3,051 | 19.97% |
|  | Senior | 3 | 4.05% | 4,309 | 28.20% |
|  | Unclassified | -- | -- | 49 | 0.32% |

**Table A2: Admissions and Cumulative GPA**

|  |  | Study Sample | ENGL 2010 Population |
|---|---|---|---|
| Admission Index (combined GPA & ACT/SAT score) | Avg. Admission Index | 117.2 | 112.5 |
|  | Std. dev. of Admission Index | 13.9 | 13.0 |
|  | Median Admission Index | 120.0 | 113.0 |
|  | Min. Admission Index | 72.0 | 62.0 |
|  | Max. Admission Index | 138.0 | 142.0 |
| Cumulative GPA | Avg. Cumulative GPA | 3.543 | 3.215 |
|  | Std. dev. of Cumulative GPA | 0.395 | 0.607 |
|  | Median Cumulative GPA | 3.576 | 3.320 |
|  | Min. Cumulative GPA | 2.368 | 0.118 |

## Appendix B: Inter-Rater Scores

**Table B1: IL VALUE Rubric Inter-Rater Reliability by Round and Rubric Category**

| Round | Cat. A | Cat. B | Cat. C | Cat. D | Cat. E |
|---|---|---|---|---|---|
| Round 1 (papers 1 – 5) | 0.685 | -0.249 | 0.152 | 0.103 | -0.273 |
| Round 2 (papers 6 – 15) | 0.444 | 0.664 | 0.317 | 0.870 | 0.225 |
| Round 3 (papers 16 – 25) | 0.434 | 0.493 | 0.589 | 0.373 | 0.334 |
| Round 4 (papers 26 – 35) | 0.394 | 0.649 | 0.479 | 0.458 | 0.037 |

**Table B2: IL VALUE Rubric Scorers Percentage of Agreement by Rubric Category**

| Category | % of 3 Scorers Agreement | % of 2 Scorers Agreement | % of No Scorers Agreement |
|---|---|---|---|
| Category A | 19.7 | 69.7 | 10.5 |
| Category B | 23.7 | 68.4 | 7.9 |
| Category C | 13.2 | 72.4 | 14.5 |
| Category D | 30.3 | 63.2 | 6.6 |
| Category E | 30.3 | 60.5 | 9.2 |

**Table B3: Synthesis Rubric Inter-Rater Reliability by Round and Rubric Category**

| Rubric Category | Round 1 | Round 2 |
|---|---|---|
| Category A | 0.512 | 0.482 |
| Category B | 0.364 | 0.646 |
| Category C | 1 | 0.738 |
| Category D | 0.563 | 0.634 |
| Category E | 0.622 | 0.671 |

[ ARTICLE ]

*Eastman, et al*
*Closing the Loop*