English 101 Evaluation & Assessment Pilot: Benchmarking Report

Carol A. Leibiger & Alan W. Aldrich

University Libraries

University of South Dakota

Author Note

The authors wish to thank the members of the University Libraries Research, Reference, and Instruction Services group for their participation in the design and conduct of the evaluation and assessment during fall of 2015. The authors also thank the instructors and students of the English 101 sections which provided the opportunity to conduct this research.

Correspondence concerning this report should be addressed to Carol Leibiger, Ph.D., University Libraries, University of South Dakota, 414 East Clark Street, Vermillion, SD, 57069. Email: C.Leibiger@usd.edu
While the university library is no longer “the heart of campus,” it continues to inhabit a variety of important roles in the intellectual and social life of higher education, one of which is as the site of teaching and learning. Evaluation and assessment afford ways for the library to demonstrate to its constituents that it is expending resources to support institutional goals through effective teaching (evaluation) and student learning (assessment) of information literacy (IL). This report documents the continuation of the pilot of a hybrid evaluation and assessment instrument in ENGL 101 (Freshman Composition) at the University of South Dakota (USD) by members of the University Libraries’ (UL) Research, Reference, and Instruction Services (RRIS) in the fall semester of 2015.

ENGL 101 is a general-education, writing-skills course required of freshman in their first semester at USD. (Exceptions are granted for students entering with prior credit for this course, e.g., for transfer students and those who have successfully passed AP English or “dual-credit” ENGL 101 while in high school). The UL supports two ENGL 101 research papers, the Research-Based Academic Argument (RBAA), a general research paper; and a second paper (either a media-analysis paper or a synthesis essay) in face-to-face library instructional sessions.

Because of the volume of ENGL 101 instruction (50 or more sections/semester on the main USD campus and a satellite campus), 12 RRIS members, 11 faculty librarians and one paraprofessional, perform instruction to support course research assignments. The UL espouses active learning, and students are brought to the library when their focus is on their assignments, and they are primed to learn. The IL Coordinator performs instructional design to ensure quality and uniformity of instruction and maximize time on task in support of active learning. Students engage in flipped instruction on research methods and tools before the library session so that RRIS members can quickly review the use of tools and searching techniques appropriate for finding resources in support of the students’ assignment. (The resources are all findable on the
Freshman English Research Guide, a LibGuide™ created for the course.) According to the instructional plan, RRIS members should spend no more than 15-20 minutes in an instructional session demonstrating the use of tools and searching strategies. This allows the students to spend the remaining time using the catalog and research databases to find books and articles in support of their research topics. RRIS members are expected to interact one-on-one with each student, answering questions and providing expert guidance.

Communication underlies most library encounters with students, and instruction shares the instrumental and relational dimensions of communication (Watzlawick, Beavin, & Jackson, 1967). The instrumental aspect of ENGL 101 IL instruction is to enable students to find the resources they need to complete their assignments. The relational dimension of the instruction is reflected in the development of a positive relationship between students on the one hand and librarians and the library on the other hand so that students will return to the UL with future research assignments.

Instructional design attempts to ensure uniformity and quality of instruction through adequate preparation and planning before teaching occurs. Evaluation and assessment seek to determine the extent to which RRIS members have provided quality instruction and facilitated learning. Evaluation and assessment provide data for individual post-instruction reflection as well as examination of programmatic effectiveness. On the basis of evaluation data, RRIS members can consider how effectively they have taught and engage in individual professional development to enhance their efficacy as instructors. Additionally, evaluation and assessment data allow the Instructional Team to suggest areas for improvement in RRIS instruction and provide appropriate professional development. Student feedback also allows the Instructional Team to gauge the effectiveness of, and revise, instructional design. Initiating a cycle of curricular planning, instruction, evaluation and assessment, reflection, and professional and curricular development and revision (the assessment cycle, which also can apply to evaluation) is
crucial for creating a culture of assessment and continuous improvement in educational institutions.

**Design of the Evaluation and Assessment Instrument**

To evaluate instruction and assess learning in ENGL 101, a hybrid evaluation and assessment instrument was developed and deployed using the Socrative™ Student Response tool’s quiz function. The instrument was developed as a collaborative project within RRIS, and it had been piloted in spring 2015 by the UL Instructional Team, the IL Coordinator and the Instructional Services Librarian (Leibiger & Aldrich, 2016). Because all ENGL 101 students must write the RBAA paper, it was chosen as the site of the evaluation/assessment pilot. The instrument was made accessible to students using the quiz function of Socrative™, a free online tool that can be used to create polls, games, quizzes, and classroom assessment techniques (CATs), and a like to the online form was placed into the Freshman English Research Guide. Students were asked to access and complete the instrument during the last five minutes of their research session. Evaluation questions consisted of four Likert-like closed-ended questions and one open-ended question. A final open-ended question assessed learning from the session, which emphasized searching using the library catalog and general-purpose research databases like *Academic Search Premier* and *ProQuest*, aligning with ACRL IL Standard 2.¹ The four closed-ended evaluation questions and the open-ended evaluation and assessment questions are listed in Table 1 below.

¹ Since 2005 the South Dakota Board of Regents has mandated IL as a system general-education requirement, and the ACRL Information Literacy Competency Standards for Higher Education provide the student learning outcomes for this requirement.
Table 1: Evaluation and Assessment Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Closed-Ended Evaluation Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The librarian presented the material effectively.</td>
</tr>
<tr>
<td>2</td>
<td>The librarian presented clear and accurate information.</td>
</tr>
<tr>
<td>3</td>
<td>The librarian answered questions competently.</td>
</tr>
<tr>
<td>4</td>
<td>Students had the opportunity to participate and/or ask questions.</td>
</tr>
<tr>
<td>Open-Ended Evaluation and Assessment Questions</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>How could the librarian have taught this session better?</td>
</tr>
<tr>
<td>6</td>
<td>What did you learn in this library session that you could pass on to fellow students or friends to help them complete this assignment better?</td>
</tr>
</tbody>
</table>

Ten of 11 RRIS members participated in the evaluation/assessment pilot in 40 sections of ENGL 101.² RRIS members culled their evaluation/assessment results from their Socrative™ accounts, copied them into a formatted Excel™ spreadsheet, and sent them to a generic e-mail account administered by a Graduate Assistant (GA) assigned to RRIS. The GA stripped all personal and identifying information from the 40 spreadsheets before making them available to the Instructional Team for analysis and benchmarking. Quantitative and qualitative analyses were conducted during the summer and fall of 2016. The Instructional Team engaged in benchmarking and further analysis of the results during the spring semester of 2017.

**Description of Quantitative Data: Evaluation Questions**

Students responded to the four quantitative questions using a Likert-like five-point scale anchored by strongly disagree = 1 and strongly agree = 5. Additional responses included disagree = 2, neither agree nor disagree = 3, and agree = 4. A total of 618 students responded. The mean scores for each question as well as the overall mean score across all sections are presented in Table 2 below.

---

² One RRIS member did not submit evaluation/assessment results, resulting in the loss of one section’s data.
Table 2: Overall Mean Scores for Evaluation Questions 1-4

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Question</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Librarian presented materials effectively</td>
<td>4.52</td>
</tr>
<tr>
<td>Q2</td>
<td>Librarian presented clear and accurate information</td>
<td>4.51</td>
</tr>
<tr>
<td>Q3</td>
<td>Librarian answered questions completely</td>
<td>4.51</td>
</tr>
<tr>
<td>Q4</td>
<td>Students could participate and ask questions</td>
<td>4.53</td>
</tr>
</tbody>
</table>

The mean scores suggest a pronounced lack of variability. To identify variability within and across sections the lowest and highest mean scores for each question are presented in Table 3 below.

Table 3: Lowest and Highest Mean Scores for Evaluation Questions 1-4

<table>
<thead>
<tr>
<th>Sections with Lowest/Highest Mean Scores for Each Question</th>
<th>Lowest mean score by section</th>
<th>Highest mean score by section</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section with the lowest mean score for Q1</td>
<td>4.19</td>
<td></td>
</tr>
<tr>
<td>Section with the highest mean score for Q1</td>
<td></td>
<td>4.90</td>
</tr>
<tr>
<td>Section with the lowest mean score for Q2</td>
<td>4.22</td>
<td></td>
</tr>
<tr>
<td>Section with the highest mean score for Q2</td>
<td></td>
<td>4.90</td>
</tr>
<tr>
<td>Section with the lowest mean score for Q3</td>
<td>4.00</td>
<td></td>
</tr>
<tr>
<td>Section with the highest mean score for Q3</td>
<td></td>
<td>4.90</td>
</tr>
<tr>
<td>Section with the lowest mean score for Q4</td>
<td>4.19</td>
<td></td>
</tr>
<tr>
<td>Section with the highest mean score for Q4</td>
<td></td>
<td>4.93</td>
</tr>
</tbody>
</table>

Description of Qualitative Data: Evaluation Question

The Instructional Team coded the data from the open-ended questions. For the evaluation question, “What could the librarian do better?” students provided 618 responses, 587 of which were usable. Responses coded as unusable were ones where students failed to answer or provided nonsensical or irrelevant comments. The numbers of responses to the open-ended evaluation and assessment questions for all sections are summarized in Table 4 below.
Table 4: Total Responses to Open-Ended Evaluation and Assessment Questions

<table>
<thead>
<tr>
<th>Spread-sheet/section no.</th>
<th>Total evaluation responses</th>
<th>Usable responses</th>
<th>Librarian positive behavior</th>
<th>Librarian positive behavior unpacked</th>
<th>Librarian recommendation</th>
<th>Librarian recommendation unpacked</th>
<th>Librarian value/affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
<td>16</td>
<td>11</td>
<td>15</td>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>13</td>
<td>10</td>
<td>13</td>
<td>4</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>29</td>
<td>14</td>
<td>15</td>
<td>12</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>14</td>
<td>14</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>11</td>
<td>9</td>
<td>13</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>13</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>11</td>
<td>9</td>
<td>12</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>11</td>
<td>7</td>
<td>11</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
<td>17</td>
<td>11</td>
<td>14</td>
<td>7</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>19</td>
<td>15</td>
<td>19</td>
<td>5</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>14</td>
<td>13</td>
<td>3</td>
<td>5</td>
<td>11</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>16</td>
<td>14</td>
<td>11</td>
<td>15</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>22</td>
<td>15</td>
<td>9</td>
<td>15</td>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>27</td>
<td>22</td>
<td>14</td>
<td>14</td>
<td>9</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>16</td>
<td>13</td>
<td>11</td>
<td>16</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>16</td>
<td>16</td>
<td>8</td>
<td>15</td>
<td>7</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>19</td>
<td>18</td>
<td>14</td>
<td>14</td>
<td>6</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>10</td>
<td>9</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>14</td>
<td>12</td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>21</td>
<td>20</td>
<td>17</td>
<td>17</td>
<td>4</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>20</td>
<td>19</td>
<td>8</td>
<td>9</td>
<td>11</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>14</td>
<td>14</td>
<td>10</td>
<td>12</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>16</td>
<td>16</td>
<td>14</td>
<td>15</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>14</td>
<td>13</td>
<td>9</td>
<td>11</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>16</td>
<td>14</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>18</td>
<td>18</td>
<td>10</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>9</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>16</td>
<td>13</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>13</td>
<td>13</td>
<td>11</td>
<td>13</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>20</td>
<td>29</td>
<td>19</td>
<td>27</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>36</td>
<td>12</td>
<td>12</td>
<td>7</td>
<td>14</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>37</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>38</td>
<td>20</td>
<td>19</td>
<td>18</td>
<td>23</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>39</td>
<td>18</td>
<td>13</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>40</td>
<td>14</td>
<td>14</td>
<td>13</td>
<td>18</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Totals</td>
<td>618</td>
<td>567</td>
<td>382</td>
<td>482</td>
<td>189</td>
<td>219</td>
<td>14</td>
</tr>
</tbody>
</table>
Students for the most part answered the evaluation question in one of two ways. First, 382 students commented positively on their respective librarian’s instruction, indicating a high degree of satisfaction. Because many students (n=100) responded with more than one positive statement, the responses were unpacked, yielding 482 positive statements. Of students expressing satisfaction with instruction, 312 (65%) answered with a “holistic positive” blanket statement of approval (“S/he did a great job.”) usually without mentioning a specific positive teaching behavior or making a recommendation for improvement. This extremely positive response appeared to validate the high overall quantitative averages mentioned above.

Sixty holistic positive responses were accompanied by the mention of other positive behaviors. These most often included effective explanations (n = 46 or 12%), helpfulness (n = 9 or 2%), and individualized attention (n = 5 or 1%). The 172 students who did not provide holistic positive comments praised their librarians for specific behaviors. Their comments gave rise to 16 themes, the most prevalent of which was effective explanations (19%). The themes derived from students’ identification of positive librarian instructional behaviors are listed in Table 5 below.

---

3 The only recommendation made in combination with a holistic positive evaluation was for longer library sessions, a suggestion beyond the control of RRIS members.
A second way that students interpreted the open-ended evaluation question was as an invitation to critique their instructor’s teaching by providing specific recommendations for improvement. A total of 189 students answered in this way, implying less than full satisfaction with instruction. Because some students (n = 20) responded with more than one recommendation, the responses were unpacked, yielding 219 unique recommendations. Additionally, 19 students responded that their librarian had taught well, yet made a recommendation for improvement in teaching. These responses were coded “generally positive” because the recommendation suggested less satisfaction with instruction than that communicated by “holistic positive” statements, which (with one exception) were not accompanied by recommendations.

The 219 student recommendations yielded 29 themes that could be grouped into larger instructional categories. These categories include: instructional management (e.g., appropriate
pacing, quality of explanations, and attending to students), explanations, self-presentation (e.g., librarian affect, appropriate volume, appropriate movement, and humor), and two recommendations that were outside the librarians’ control of the instructional context (sessions should be longer). Most students (n = 126 or 56%) made recommendations relating to instructional management; of these, 40% (n = 87) recommended that librarians pace themselves better. In the next highest category, students valued detailed or effective explanations (n = 66 or 29%). In the category of self-presentation, 13 students recommended that librarians display more positive affect (6%), and 7 students suggested that librarians should speak with appropriate volume (3%). The categories and themes with their raw numbers and percentages are presented in Table 6 below.
<table>
<thead>
<tr>
<th>Student Recommendations</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instruction management</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>use appropriate pacing</td>
<td>87</td>
<td>40%</td>
</tr>
<tr>
<td>provide more individual attention to students</td>
<td>13</td>
<td>6%</td>
</tr>
<tr>
<td>provide more interactive learning</td>
<td>6</td>
<td>3%</td>
</tr>
<tr>
<td>be attentive to class</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>provide more interesting instruction</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>offer more opportunities for questions</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>answer questions effectively</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>offer more opportunities for hands-on work</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>be better prepared</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>manage questions better</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>model effective searching</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>provide more active learning</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>restate/repeat instruction</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>126</td>
<td>56%</td>
</tr>
<tr>
<td><strong>Explanations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>provide more detailed explanations</td>
<td>14</td>
<td>6%</td>
</tr>
<tr>
<td>provide more detailed explanations (databases)</td>
<td>16</td>
<td>7%</td>
</tr>
<tr>
<td>provide more detailed explanations (library resources)</td>
<td>6</td>
<td>3%</td>
</tr>
<tr>
<td>provide more detailed explanations (library web site)</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>provide more effective explanations</td>
<td>19</td>
<td>9%</td>
</tr>
<tr>
<td>provide more effective explanations (databases)</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>provide more effective explanations (library resources)</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>provide more focused examples</td>
<td>7</td>
<td>3%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>66</td>
<td>29%</td>
</tr>
<tr>
<td><strong>Self-presentation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>provide more positive affect</td>
<td>13</td>
<td>6%</td>
</tr>
<tr>
<td>use appropriate volume</td>
<td>7</td>
<td>3%</td>
</tr>
<tr>
<td>decrease movement</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td>make eye contact</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>use humor</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>speak more clearly</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>25</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Beyond instructor control</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>provide longer sessions</td>
<td>2</td>
<td>1%</td>
</tr>
<tr>
<td><strong>Total recommendations</strong></td>
<td>219</td>
<td>96%</td>
</tr>
</tbody>
</table>
Finally, 14 student responses noted affect associated with librarians’ instruction. Nine responses align with levels of affect associated with learning in Bloom’s Taxonomy: Affective Domain (Clark, 2004). One student (7%) noted an awareness of having learned, which aligned with level 1 (awareness, attention). Additionally, four students (29%) indicated satisfaction with instruction, corresponding to level 2 (responding, satisfaction in responding). Lastly, four students (29%) noted confidence as a result of librarians’ instruction; confidence reflects the fifth (highest) level in Bloom’s affective taxonomy (internalizing values, self-reliance). Responses reflecting affect and values are displayed in Table 7 below.

<table>
<thead>
<tr>
<th>Affect/Value</th>
<th>n</th>
<th>%</th>
<th>Bloom’s Taxonomy: Affective Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>confidence</td>
<td>4</td>
<td>29%</td>
<td>Level 5 (self-reliance, self-efficacy)</td>
</tr>
<tr>
<td>gratitude</td>
<td>4</td>
<td>29%</td>
<td></td>
</tr>
<tr>
<td>lack of confidence</td>
<td>1</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>satisfaction</td>
<td>4</td>
<td>29%</td>
<td>Level 2 (responding)</td>
</tr>
<tr>
<td>surprise at learning</td>
<td>1</td>
<td>7%</td>
<td>Level 1 (awareness)</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>101%</td>
<td></td>
</tr>
</tbody>
</table>

**Description of Qualitative Data: Assessment Question**

The five evaluation questions were followed by an open-ended assessment question: “What did you learn in this library session that you could pass on to fellow students or friends to help them complete this assignment better?” Of 618 student responses, 539 were useable. Responses coded as unusable were ones where students failed to answer or provided nonsensical or irrelevant comments.

Students provided 435 responses specifying skills they had learned or improved upon during the library session. Fifty-nine responses mentioned two or more skills; once unpacked, there were 499 responses identifying skills derived from the library instruction. Since the library sessions taught catalog and database searching in support of a research paper, it was gratifying to
see that 307 students (62%) stated that they had learned searching. An additional 36 students (7%) indicated that they had improved in searching. A total of 343 students (68%) identified searching, the skill that RRIS members had taught, as the learning derived from instruction.

Some students (n = 129 or 26%) identified learning of specific aspects of information finding: navigating the library web site, accessing library resources, narrowing the search, evaluating resources, and managing information), or an activity not directly related to searching, using information responsibly, that aligns with ACRL IL Competency Standard 5 (ACRL, 2000). A large group of students (n = 85 or 17%) indicated that they had learned to navigate the library web site. The skills identified by students are listed in Table 8 below.

<table>
<thead>
<tr>
<th>Skill</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>searching</td>
<td>307</td>
<td>62%</td>
</tr>
<tr>
<td>improvement in searching</td>
<td>36</td>
<td>7%</td>
</tr>
<tr>
<td>navigating the library web site</td>
<td>85</td>
<td>17%</td>
</tr>
<tr>
<td>narrowing the search</td>
<td>38</td>
<td>8%</td>
</tr>
<tr>
<td>managing information</td>
<td>26</td>
<td>5%</td>
</tr>
<tr>
<td>evaluating resources</td>
<td>3</td>
<td>1%</td>
</tr>
<tr>
<td>navigating library resources</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>accessing library resources</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>using information responsibly</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>499</td>
<td>100%</td>
</tr>
</tbody>
</table>

Students also mentioned tools that they had learned to use during the library instruction; they provided 421 responses naming resources, which, when unpacked, gave rise to 450 unique data points. Students overwhelmingly mentioned having used library resources (n = 267, or 59%), research databases (n = 110, or 24%), and the library catalog (n = 19, or 4%). In total 396 students (88%) stated that they had learned to use at least one of the resources that had been taught during the library session. The next highest group students (n = 25, or 6%) indicated that they had learned to use the course LibGuide™. Skills mentioned by students are listed in Table 9.
below.

<table>
<thead>
<tr>
<th>Table 9: Student Learning of Tools from the Open-Ended Assessment Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource</td>
</tr>
<tr>
<td>Library research resources, databases, catalog taught (total)</td>
</tr>
<tr>
<td>library resources</td>
</tr>
<tr>
<td>research databases</td>
</tr>
<tr>
<td>library catalog</td>
</tr>
<tr>
<td><strong>Other resources (total)</strong></td>
</tr>
<tr>
<td>LibGuide™</td>
</tr>
<tr>
<td>interlibrary loan</td>
</tr>
<tr>
<td>SFX</td>
</tr>
<tr>
<td>reference service</td>
</tr>
<tr>
<td>library web site</td>
</tr>
<tr>
<td>Discovery layer</td>
</tr>
<tr>
<td>OWL</td>
</tr>
<tr>
<td>government documents</td>
</tr>
<tr>
<td>university web site</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

Finally, 34 students identified changes in their affect or values as a result of library instruction. The greatest affective change noted was an increase in confidence in searching (n = 8 or 24%). Additionally, 23 students indicated that they had come to value library resources (n = 21 or 62%), library instruction (n = 1 or 3%), or librarians (n = 1 or 3%). If confidence is understood as self-efficacy or self-reliance, students who reported experiencing increased confidence can be placed on the highest (fifth) level of Bloom’s Taxonomy of the Affective Domain (Clark, 2004). Students who identified an increase in value can be placed on the third level of the taxonomy (valuing learning). Students’ reported changes in values and affect are listed in Table 10 below.
Table 10: Student Affect/Values from the Open-Ended Assessment Question

<table>
<thead>
<tr>
<th>Affect/values associated with learning</th>
<th>n</th>
<th>%</th>
<th>Bloom’s Taxonomy: Affective Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>confidence</td>
<td>8</td>
<td>24%</td>
<td>Level 5 (self-reliance, self-efficacy)</td>
</tr>
<tr>
<td>gratitude</td>
<td>2</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>perseverance</td>
<td>1</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td><strong>Valuing (total)</strong></td>
<td>23</td>
<td>68%</td>
<td>Level 3 (valuing)</td>
</tr>
<tr>
<td>valuing librarians</td>
<td>1</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>valuing library instruction</td>
<td>1</td>
<td>3%</td>
<td></td>
</tr>
<tr>
<td>valuing library resources</td>
<td>21</td>
<td>62%</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>34</td>
<td>101%</td>
<td></td>
</tr>
</tbody>
</table>

**Analysis**

The mean scores for the quantitative questions 1-4 were very high and suggested that RRIS members were providing excellent instruction. However, the amount and type of information provided by the students answering the open-ended evaluation question suggested there was room for improvement in instruction. In seeking to manage these seemingly contradictory sets of data, the Instructional Team had to manage two issues in the analysis: The first concern was how best to analyze the quantitative responses to questions 1-4 given the high mean scores and low degree of variability between the low and high mean scores. The second concern was how to best relate the two sets of data in order to identify areas of strength and weakness in RRIS instruction by deploying methods of hybrid analysis.

**Stratification of Data**

As indicated in Table 2, the lowest mean score for questions 1-4 across all sections was M=4.00 (Q3) and the highest mean score was M=4.93 (Q4). Because all mean scores were located within this narrow bandwidth, a quartile analysis was considered. The researchers noticed natural break points in both the low and high mean scores for each question across the 40 sections such that they could meaningfully compare the lowest and highest 20% (quintiles) of mean scores for each question in every section including ties. This resulted in the eight sections...
with the lowest and highest mean scores undergoing analysis for Q1-3 and the nine sections with the lowest and highest mean scores for Q4, including ties. See Table 11 below.

<table>
<thead>
<tr>
<th>Q1 8 Lowest Average Scores</th>
<th>Q2 8 Lowest Average Scores</th>
<th>Q3 8 Lowest Average Scores</th>
<th>Q4 8 Lowest Average Scores &amp; Ties</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM-01</td>
<td>BM-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-03</td>
<td>BM-06</td>
<td>BM-06</td>
<td>BM-06</td>
</tr>
<tr>
<td>BM-10</td>
<td>BM-10</td>
<td>BM-10</td>
<td>BM-10</td>
</tr>
<tr>
<td>BM-17</td>
<td>BM-21</td>
<td>BM-22</td>
<td>BM-22</td>
</tr>
<tr>
<td>BM-26</td>
<td>BM-27</td>
<td>BM-28</td>
<td>BM-28</td>
</tr>
<tr>
<td>BM-29</td>
<td>BM-31</td>
<td>BM-32</td>
<td>BM-32</td>
</tr>
<tr>
<td>BM-32</td>
<td>BM-35</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q1 8 Highest Average Scores</th>
<th>Q2 8 Highest Average Scores</th>
<th>Q3 8 Highest Average Scores</th>
<th>Q4 8 Highest Average Scores &amp; Ties</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM-01</td>
<td>BM-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-05</td>
<td>BM-05</td>
<td>BM-05</td>
<td>BM-05</td>
</tr>
<tr>
<td>BM-07</td>
<td>BM-07</td>
<td>BM-07</td>
<td>BM-15</td>
</tr>
<tr>
<td>BM-15</td>
<td>BM-18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-20</td>
<td>BM-21</td>
<td>BM-23</td>
<td></td>
</tr>
<tr>
<td>BM-25</td>
<td>BM-25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM-33</td>
<td>BM-34</td>
<td>BM-34</td>
<td>BM-35</td>
</tr>
<tr>
<td>BM-34</td>
<td>BM-35</td>
<td>BM-37</td>
<td>BM-37</td>
</tr>
<tr>
<td>BM-38</td>
<td>BM-38</td>
<td>BM-38</td>
<td>BM-38</td>
</tr>
<tr>
<td>BM-39</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 12 below identifies the number of times each section appeared in the low mean score quintiles across questions 1-4. A total of 15 sections were present in the low mean score quintiles and five sections (33%) appeared in either three or four of the low mean score quintiles. Five sections appeared in two of the four low mean score quintiles (33%). Thus a majority of sections (66%) appeared in at least two or more of the four low mean score quintiles.

<table>
<thead>
<tr>
<th>Section</th>
<th>Occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM-10</td>
<td>4</td>
</tr>
<tr>
<td>BM-32</td>
<td>4</td>
</tr>
<tr>
<td>BM-06</td>
<td>3</td>
</tr>
<tr>
<td>BM-26</td>
<td>3</td>
</tr>
<tr>
<td>BM-28</td>
<td>3</td>
</tr>
<tr>
<td>BM-02</td>
<td>2</td>
</tr>
<tr>
<td>BM-03</td>
<td>2</td>
</tr>
<tr>
<td>BM-22</td>
<td>2</td>
</tr>
<tr>
<td>BM-27</td>
<td>2</td>
</tr>
<tr>
<td>BM-31</td>
<td>2</td>
</tr>
<tr>
<td>BM-01</td>
<td>1</td>
</tr>
<tr>
<td>BM-17</td>
<td>1</td>
</tr>
<tr>
<td>BM-21</td>
<td>1</td>
</tr>
<tr>
<td>BM-29</td>
<td>1</td>
</tr>
<tr>
<td>BM-35</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 13 below identifies the number of times each section appeared in the high mean score quintiles across questions 1-4. A total of 17 sections were present in the high mean score quintiles and six sections (35%) appeared in either three or four of the high mean score quintiles. Two sections appeared in two of the four high mean score quintiles (12%). Thus just under a majority of sections (47%) appeared in at least two or more of the four high mean score quintiles.
To test the hypothesis that the lowest and highest 20% of the sections had mean scores different from each other beyond chance, One Way ANOVA with Independent Samples k=2 was performed for the lowest and highest quintile mean scores for each question. The quintiles were statistically different. For question 1, the low mean score quintile (M = 4.25) was significantly different from the high mean score quintile (M = 4.77), F(1,14) 308.71 p < .0001. For question 2, the low mean score quintile (M = 4.31) was significantly different from the high mean score quintile (M = 4.72), F(1,14) 184.04 p < .0001. For question 3, the low mean score quintile (M = 4.20) was significantly different from the high mean score quintile (M = 4.79), F(1,14) 172.07 p < .0001. For question 4, the low mean score quintile (M = 4.31) was significantly different from the high mean score quintile (M = 4.84), F(1,14) 184.11 p < .0001.

In order to perform statistical analysis comparing qualitative with quantitative data, the qualitative data has to be expressed in numerical terms. To this end the qualitative comments were coded into discrete categories (see Tables 4-10 above), after which the number of responses

| Table 13: Number of Occurrences of Sections in the High Mean Score Quintiles for Q1-4 |
|-----------------------------------|------|
| BM-38                             | 4    |
| BM-40                             | 4    |
| BM-05                             | 3    |
| BM-07                             | 3    |
| BM-34                             | 3    |
| BM-37                             | 3    |
| BM-25                             | 2    |
| BM-39                             | 2    |
| BM-01                             | 1    |
| BM-04                             | 1    |
| BM-15                             | 1    |
| BM-18                             | 1    |
| BM-20                             | 1    |
| BM-21                             | 1    |
| BM-23                             | 1    |
| BM-33                             | 1    |
for each category were tallied across the low and high quintiles for Q1-4. The number of responses for each category was expressed as a proportion (see Tables 4-10 above for all percentages). A series of Z tests was conducted to test the relationship of each proportion to the different quintiles of each quantitative question.

Results

Evaluation Comments Related to Q1: Librarian Presented Materials Effectively

The following categories were statistically significant in relationship to question 1: Just as with the overall dataset, instructors in the high score quintile on Q1 were seen as being significantly more helpful (.07) than instructors in the low score quintile on Q1 (.01), $Z = -2.20$, $p < .05$.

Many of the comments made by students in their responses to the evaluation question included specific suggestions or recommendations for improvement. Instructors in the low score quintile on Q1 received significantly more comments about the need to provide more detailed explanations (.07) than instructors in the high score quintile on Q1 (0.0), $Z = 2.08$, $p < .05$.

Instructors in the low score quintile on Q1 also received significantly more comments about the need for displaying positive affect (.11) than instructors in the high score quintile on Q1 (0.0), $Z = 2.59$, $p < .05$.

Evaluation Comments Related to Q2: Librarian Presented Clear and Accurate Information

Instructors in the low score quintile on Q2 received significantly more comments about the need for appropriate pacing (.40) than did instructors in the high score quintile on Q2 (.14), $Z = 2.72$, $p < .05$.

Evaluation Comments Related to Q3: Librarian Answered Questions Completely

Mirroring the results for the overall dataset, instructors in the high score quintile on Q3 were seen as being significantly more helpful (.10) than instructors in the low score quintile on
Q3 (.01), Z = -2.72, p < .05. Conversely, instructors in the low score quintile on Q3 received significantly more holistic positive comments (.75) than instructors in the high score quintile on Q3 (.58): Z = 2.43, p < .05.

**Evaluation Comments Related to Q4: Students Could Participate and Ask Questions**

Instructors in the high score quintile on Q4 were seen as being significantly more helpful (.07) than instructors in the low score quintile on Q4 (.01), Z = -2.27, p < .05.

**Analysis of Assessment**

The final open-ended question was designed to assess student learning through self-identification of skills and tools learned. Just as with the open-ended evaluation responses the skills identified by the respondents were organized into inductive categories, and Z tests were done across the low and high score quintiles testing for relationships between each quantitative question and related skills learned by the students.

**Assessment Comments Related to Q1: Librarian Presented Materials Effectively**

*Narrowing the search* was a skill identified by significantly more students in the high score quintile on Q1 (.16) compared to students in the low score quintile on Q1 (.07), Z = -1.95, p < .05. At the same time, *Navigating the library web site* was a skill reported by significantly more students in the low score quintile on Q1 (.22) compared to the high score quintile on Q1 (.12), Z = 1.97, p < .05.

**Assessment Comments Related to Q2: Librarian Presented Clear and Accurate Information**

The assessment results related to question 2 are identical to the results for question 1. *Narrowing the search* was a skill identified by significantly more students in the high score quintile on Q2 (.16) compared to students in the low score quintile on Q2 (.06), Z = -2.33, p < .05. At the same time, *Navigating the library web site* was a skill reported by significantly
more students in the low score quintile on Q2 (.26) compared to students in the high score quintile on Q2 (.10), Z = 3.26, p < .05.

Assessment Comments Related to Q3: Librarian Answered Questions Completely

Improvement in searching was a skill identified by significantly more students in the low score quintile on Q3 (.11) compared to students in the high score quintile on Q3 (.02), Z = 2.37, p < .05.

Assessment Comments Related to Q4: Students Could Participate and Ask Questions

There were no significant differences for skills reported by students in either the low or high score quintiles on Q4.

Discussion

The uniformity of the consistently high mean scores for questions 1-4 across all sections caused the researchers to stratify the data into quintiles with the lowest and highest quintiles subjected to further analysis. There are several possible explanations for these results. Students’ ability to discriminate along a Likert-like range appears to be limited as they do not have a clear sense of ranges or degrees in terms of rating others. This might be due to students’ being accustomed to receiving generally high or inflated marks on their academic work, such that they are conditioned to thinking in the same terms when evaluating instruction. This consistency in students’ responses to the quantitative evaluation questions 1-4 suggests the importance of including open-ended responses in studies such as this one. Open-ended response questions can help in understanding quantitative data arrayed along a very narrow bandwidth.

The breakdown of the quintiles presented in Tables 12 and 13 above show relative consistency in discriminating between low performing instructors and high performing instructors. There were several sections that appeared once in both the low and high mean score quintiles. Also there were many more sections that appeared once in the high mean score quintiles (nine sections) than those that appeared once in the low mean score quintiles (five sections).
Appearing once in either a low or high mean score quintile could be due to the instructor’s having an “off” day or an “on” day in an area relative to a particular category.

Instructors who can be identified as either “low performers” or “high performers” in this analysis are those whose sections consistently appear in either three or four of the low or high quintiles, respectively. The low-performing instructors should consider reflecting upon and working towards improving their teaching using data from the evaluations of their sections and the conclusions of this study. Conversely, high-performing instructors should also reflect upon their teaching with the goal of identifying what it is that they are doing well so they can (a) reinforce those behaviors or techniques and (b) share their findings with the other faculty members, including those who are in the middle quintiles where the teaching is adequate to good but still has room for improvement.

**Evaluation Results-Quantitative and Qualitative Data in Dialog**

The high means in the quantitative results for closed-ended questions 1 through 4 suggest on the surface that the library faculty is providing good instruction that is valued by the students across all of the English 101 sections. The qualitative data indicate that students are not as uniformly satisfied with instruction as the quantitative data suggest. The Instructional Team chose to put the qualitative data in conversation with the quantitative data to get a more nuanced analysis that would help to understand this discrepancy and tease out differences in instruction among the low and the high performers.

Based on the results of this study, it is obvious that students attend to the instruction they receive from RRIS. Given the constructivist design of the library IL instruction to include student active learning facilitated by RRIS members as “guides on the side”, it is not a surprise that instructors scoring high on the quantitative questions are perceived as being significantly more helpful than instructors with low scores. This helpfulness is accentuated by those instructors who make time to work individually with each student during the IL session.
Helpfulness was significantly correlated with the high performers in relation to questions 1, 3, and 4.

Question 1 asked whether the librarian presented materials effectively or not. Instructors with low scores on this question were also perceived as needing to provide more detailed or complete explanations by their students. Zero comments about the need for more detailed information were made about instructors with high scores on question 1. The students showed through their comments how closely they attend to the quality of information being shared during IL sessions. This data reinforces the importance of planning so that the short (15-20 minute) demonstration of resources and search strategies at the beginning of library instructional sessions addresses student assignments with adequate content and detail while leaving sufficient time for student active learning focused on searching.

Another interesting finding in response to Q1 was the highlighting of instructor’s affect by the students. Comments such as “I think that the librarian could put a little more enthusiasm into it” or “Be more engaging with the students” show the students attending as much to the relational component of teaching/learning as they do to the instrumental or informational dimensions of the demonstration of tools/searching strategies and the facilitation of active learning noted above. Conversely, helpfulness, which has both instrumental and relational aspects, is associated with those instructors who received high scores.

It is important to recognize and manage the affective dimensions of teaching especially in the case of single or “one-shot” IL sessions, since affect can support or undermine cognitive activities like learning. The instructors of record for any given course or section have an entire semester in which they can develop relationships; in such a case, the impact of a single “off” day in the classroom can be minimized. Library instructors, however, do not have the luxury of developing a relationship across repeated interactions with students. They must therefore become mindful of, and manage, the relational components of instruction to the same extent that they
manage the instrumental aspects of instruction, if they are to experience success with their students.

Another common theme across many of the questions was the abundant use of holistic positive comments by students in response to the open-ended evaluation question. While all instructors received holistic positive comments, instructors with low scores on question 3 received significantly more holistic positive comments than the instructors with high scores. In questions 1, 2, and 4, the percentage of holistic positive comments received by instructors with low scores, though not statistically significant, was almost constantly 10% higher than the instructors with high scores as illustrated in Table 14 below.

<table>
<thead>
<tr>
<th>Question</th>
<th>Holistic positive % low scores</th>
<th>Holistic positive % high scores</th>
<th>Z test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.67</td>
<td>0.58</td>
<td>Z = 1.33 ns</td>
</tr>
<tr>
<td>2</td>
<td>0.7</td>
<td>0.6</td>
<td>Z = 1.48 ns</td>
</tr>
<tr>
<td>3</td>
<td>0.75</td>
<td>0.58</td>
<td>Z = 2.43, p &lt; .05</td>
</tr>
<tr>
<td>4</td>
<td>0.74</td>
<td>0.64</td>
<td>Z = 1.56 ns</td>
</tr>
</tbody>
</table>

The researchers believe that the prevalence of more holistic positive comments for instructors scoring low on the quantitative questions is a form of relational repair work on the part of the students. People in South Dakota specifically, and in the Midwest generally, engage in a form of politeness colloquially referred to as “Midwestern Nice”, a form of hyperpoliteness that is manifested in a disinclination to provide negative responses even when they are warranted. When students assign an instructor a lower rating on the quantitative scale, students seem to compensate for this by providing a holistic positive response to the open-ended evaluation question which, through its positive tone, acts to offset the lower numerical rating. This functions to minimize the perceived potential damage of a lower quantitative rating, helping to repair the relationship. An alternative explanation would be that the generic nature of the
holistic positive comment is a way to avoid committing to a specific position or is done out of mental laziness.

The language of the open-ended evaluation question “what could the librarian do better” also plays a role in generating these results, especially in the context of Midwestern nice. The question as phrased could be understood as implying that the instructor provided less than satisfactory instruction. However, students could interpret the question as providing an opportunity to offer helpful advice, which is permissible in accordance with Midwestern Nice. It is beyond the scope of this study to delve into the reasons why students responded as they did.

The fact that instructors with low scores received an average of 10% more holistic positive comments than those instructors with high scores suggests the following metric be employed by faculty when interpreting holistic positive comments. Proportions of holistic positive comments up to about 60% seem to reflect high-quality instruction. Instructors who receive holistic positive comments at a proportion higher than 60% tend to be located in the low mean score quintile. Thus, they should consider examining their quantitative data and open-ended evaluation comments to determine whether there might be teaching issues that need to be addressed.

The one area of significance that was related with question 2 was on pacing. Students who rated instructors low on question 2 made significantly more comments identifying pacing as an instructional management issue. Comments included slowing the pace of presentation and allowing more time for students to practice research skills. Students were also sensitive to each student receiving relatively equal amount of individual attention on the part of the librarian. Pacing is an issue central to successful teaching especially within a constructive paradigm that espouses interactive learning.
Assessment Results-Quantitative and Qualitative Data in Dialog

Several interesting and worrying results emerge when the results of the open-ended assessment question are put into dialog with the quantitative questions. Questions 1 and 2 had essentially identical results in terms of the statistically significant qualitative responses. Students who rated their instructors high on presenting material effectively and providing clear and accurate information indicated in significant numbers that they learned how to narrow their search as a key skill outcome.

Conversely, a large group of students (n = 85 or 17%) indicated that they had learned to navigate the library web site. Moreover, students who rated their instructors low on questions 1 and 2 indicated in significant numbers that they learned how to navigate the library website. This suggests several weaknesses in library instruction that should be addressed in future revision of instructional design.

It appears that navigating the library web site might have distracted students from the intended focus of the library session, which was searching for scholarly resources. Additionally, it seems that students receiving instruction in the low-rated sections might not have learned IL skills, since navigating the library website is a precursor to developing IL skills, but is not a skill of the order of successfully searching using library resources makes available. Students who claim learning in terms of narrowing their searches are already navigating the library website successfully and have moved from mere locational or access knowledge (the lowest level in Bloom’s taxonomy) to performing higher order IL skills, which correspond to the fourth (analyze) or fifth (evaluate) level of Bloom’s taxonomy (Armstrong, 2017).

While not appearing at the level of significance some students reported that they learned how to use the course LibGuide™. Akin to navigating the library website, which like a LibGuide™ is an institution-specific receptacle for library resources rather than a tool itself, this also suggests students are learning where information can be found but not how to manage the
finding of information and the focusing of one’s research. This aligns with concerns expressed by ENGL 101 faculty, who worry that students are unable to locate the research databases or library catalog on the library web site if they use a LibGuide™ during library instruction.

Library instructional design might have to be revised so that instructors focus on the library resources and their use, rather than on the library web page or the research guides. Alternatively, the databases might need to be located where students can access them more easily and intuitively, so that students’ focus is on their use of databases rather than navigating the library web page or a research guide.

A significant number of students in the low mean score quintile for question 3 cited improvement in searching. This is puzzling given the results of narrowing the search reported by students in the high mean score quintiles for questions 1 and 2. It could be understood as the result of student motivation, which can support learning even when instruction is less than optimal. This finding requires further research.

**Limitations of the Study**

This study has several limitations. First, it focused on a specific group of instructors, members of RRIS at the University Libraries and their students. Because the goal was to examine the teaching of all RRIS members, and the number of library instructors is low (N = 11), randomization of the sample was not possible. An omission of a group member becomes an issue for internal validity, as happened in this study. Ten of the 11 RRIS members contributed evaluation/assessment data across all ENGL 101 sections that they taught for the RBAA assignment. Thus, we cannot generalize to the entire RRIS group regarding teaching and learning.

Second, the results are also challenged by a personnel change. The Science Librarian, who performed more instruction than any other RRIS members except for the Instructional Team members, has since left USD. The Science Librarian was also perceived as a competent
instructor who might have had scores in the high quintile, thus contributing to the overall high mean scores.

Third, the Instructional Team contributed seven sections of data to this study, amounting to approximately one sixth of the total sections taught. Since they are seasoned instructors, they might have skewed the results higher.

Fourth, there are limitations inherent in self-reporting, including its well-known tenuous link to cognitive processes. In the case of the quantitative questions, students are forced to match their perceptions with preset response choices. The open-ended questions appear to allow students to express their experience. However, the extent to which written responses are genuinely reflective of student experiences is subject to question.

While acknowledging these limitations, the researchers also recognize the benefits afforded by this hybrid instrument. The instrument is efficient, utilizing minimal class time while organizing the data set for easy retrieval. It allows the collection of rich natural-language data that can be quickly and easily coded. The instrument is designed to reflect the instrumental and relational aspects of instruction. The open-ended questions as phrased encourage students to reflect on the teaching and learning they have experienced. The last (assessment) open-ended question was strategically designed to encourage students to think about their learning in a positive way, allowing them to exit instruction recognizing what they had learned and feeling good about it. Good assessment always includes learning, and meta-cognitive reflection enables students to realize that the library can serve their instrumental and relational needs.

**Reassessing the internal validity of this study**

The most important concern of the limitations described above was the threat to the internal validity caused by the omission of one instructor’s data. Without violating the anonymity of the RRIS instructors, the researchers were able to determine that one person who taught one
section did not submit data. Since there were several instructors that taught only single sessions anonymity was preserved.

A detailed secondary analysis was done to test the reliability of the data already gathered. It is hard to determine the effect that missing quantitative data would have had upon the study. However, given the large number of students responding to questions 1-4 (N = 618), the presence or absence of scores from one additional section would not have made much difference in the overall mean scores for each question. It is possible, however, to do a hypothetical analysis of the impact of the one section’s worth of qualitative data.

In the initial analysis the quantitative responses to questions 1-4 were stratified for analysis using the lowest and highest 20% of mean scores. Additionally, the qualitative responses were coded and the proportion for each category was calculated. Z tests were done for both the low and high quintile conditions to examine the relationship between each quantitative question and the proportion of qualitative responses for each category.

The secondary analysis tested two questions. First, what would happen to the tests of significance if the missing instructor had received student comments on each opened question category in the low mean score quintile? The researchers added one to the count (n + 1) for every comment category in the low quintile condition then reran the Z score analysis for each category as it relates to the quantitative questions 1-4.

The presence of the hypothetical data in the low quintile condition had the following impacts. For the librarian positive behaviors correlated with question one, an instructor in the low mean score quintile scored as being helpful would have rendered that finding non-significant, Z = -1.63, ns. The odds of a person with a mean score in the low quintile being judged helpful is unlikely. For the librarian positive behaviors correlated with question two, an additional instantiation of helpfulness in the low mean score quintile would have rendered that
finding non-significant, \( Z = -1.65 \), ns. Again, the chance that a person in the low quintile would have been perceived as helpful is unlikely given what was observed in the data set.

Second, what would happen to the tests of significance if the missing instructor had received student comments on each opened question category in the high mean score quintile? The researchers added one to the count (\( n + 1 \)) for every comment category in the high quintile condition, then reran the \( Z \) score analysis for each category as it relates to the quantitative questions 1-4.

The potential impact was greater when the hypothetical scores were added to each category in the high mean score quintile. In the recommendations for librarians related to question 1, needing to provide more detailed explanations would have become non-significant, \( Z = .60 \), ns. The recommendation that more positive affect be displayed by the instructor would have also been rendered non-significant had the missing instructor received this comment in the high quintile. An instructor performing at a high level would be less likely to receive these kinds of recommendations.

The rest of the changes that would have resulted in adding an additional result for each of the categories in the high mean score quintile are found within the librarian positive behaviors for questions 2 and 3. Another instantiation of providing individual attention observed as a positive behavior in the high mean score quintile for question 2 would have yielded significance, \( Z = -2.06, p < .05 \). For question 3, another instantiation of providing answers to questions in the high quintile would have yielded significance, \( Z = -2.04, p < .05 \). An increase in providing individual attention in question 3 in the high quintile would also have yielded significance, \( Z = -2.04, p < .05 \). In each of these cases for question 2 and 3 the probability that someone in the high mean score quintile received comments recognizing these positive attributes is very possible as these attributes are most closely associated with instructors in the high quintile.
In each case where significant findings were rendered non-significant or vice versa, the hypothetical results are consistent with those found in the overall data set on the relationship between the quantitative scores on questions 1-4 as related to each of the categories derived from the coding of the qualitative evaluation question. This increases the researchers’ confidence in the reliability of the study as reported in spite of the missing data. While this study doesn’t address the teaching of every RRIS members, the findings clearly identify both desirable teaching performance and areas for improvement.\footnote{Additional secondary analysis was performed with the assessment data as well. The results are consistent with the findings reported above.}

**Recommendations**

The researchers make the following specific recommendations based upon the results of this study. The hybrid analysis demonstrated that students are very attentive to instruction and that they value both the demonstration and interactive parts of library sessions. Planning is crucial for effective instruction. Planning includes pacing, i.e., appropriate speed of delivery and allocation of time between frontal and interactive teaching. Methods include ways of delivering content (effective, detailed explanations using relevant examples) and interacting with students to support active learning. Instructors need to ensure that they interact equally with each student. Content of instruction should be focused on higher-order skills and library research resources, i.e., databases, rather than on receptacles of these resources, i.e., the library web site or a LibGuide™.

Students also attend to relational aspects of teaching. Positive or negative affect is known to play important roles in learning. If students are to develop positive relationships with the library, instructors need to model enthusiasm for IL, teaching, and interacting with students. Instructors also need to be cognizant of how their behavior affects students, their learning, and the development of a relationship with the library.
Given the context of South Dakota, all instructors should expect to receive holistic positive comments in response to the open-ended evaluation question. However, it is important to take a nuanced view of these extremely positive comments. Holistic positive comments at a proportion over 60% are correlated with lower performance in teaching. Instructors receiving holistic positive comments at a proportion higher than 60% should examine their quantitative data and open-ended evaluation comments to determine what teaching issues might need to be addressed.

The list of student-generated teaching behaviors identified in this study can serve in the development of peer-evaluation metrics. In the context of Responsibility Center Management, students are viewed as customers. Incorporating student feedback in generating faculty evaluation criteria recognizes this important student role and also performs the Freirean function of empowering students within the educational context.

**Conclusion**

A hybrid evaluation and assessment instrument was deployed in ENGL 101 RBAA instructional sessions during the fall of 2015. The initial quantitative and qualitative findings were very positive. However, the content and number of student recommendations suggested the need for further, more nuanced analysis. This was conducted by putting the qualitative and quantitative data into dialogue, which allowed the identification of both effective and less than effective teaching behaviors.

The overly positive findings could have been due to Midwestern Nice, which needs to be managed because it can skew evaluation results, masking potential instructional issues. Hyper-politeness like Midwestern Nice manifests itself in higher quantitative mean scores and in overly positive and under-informative qualitative results. When data is put into dialog as part of the analysis, meaningful relationships between the qualitative and quantitative data can be
uncovered. These relationships allow instructors to understand holistic positive results as either indicating good instruction or issues with instruction.

This study has identified specific teaching behaviors that enhance learning within the constructivist approach used in IL instruction by RRIS. Additionally, behaviors that impair student learning were identified. Incorporating open-ended responses gives students a clear voice in the evaluation and assessment cycles. Student-generated positive and negative teaching behaviors provide valuable input to the initiation of peer review of teaching, the next step in generating a culture of evaluation, assessment, and continuous improvement.

Future directions include repeating and testing benchmarking hypotheses generated in this study with another iteration in which all RRIS members participate. RRIS members should use these findings to examine and understand their teaching, with the goal of improvement. On the basis of further iteration of this study, instruction can be revised and professional development opportunities initiated. Finally, further study is needed to describe and understand the role of hyper-politeness in evaluation and assessment.
References


