

# Citation Analysis as a Tool to Measure the Impact of Individual Research Consultations

Thomas L. Reinsfelder

This study sought to determine the degree to which individual research consultations with a librarian can improve the work of undergraduate students. Citation analysis was used to evaluate the quality of sources selected on draft papers before meeting with a librarian and on final papers after meeting with a librarian. The rating scale presented here offers guidelines for measuring the quality of sources used by students. Findings of this research begin to provide some quantitative evidence demonstrating the positive impact of individual research consultations.



It is not unusual for librarians to interact with undergraduates during the beginning stages of the research process.

Through relatively brief instruction sessions and reference desk encounters, we strive to guide students toward the use of information sources that are both high quality and appropriate for a college-level course assignment. However, in later stages, students do not seem to seek out assistance as readily, resulting in less interaction between student and librarian.

Discussions with faculty and students confirmed suspicions that many students are using information sources considered to be inadequate for scholarly or professional work. Information may be from a source whose credibility is questionable or may be completely undocumented. Due to an interest in seeing more appropriate resources used in course projects, faculty/librarian partnerships were

sought out in an effort to work more closely with students. One-on-one meetings, or research consultations, are one way to increase the level of interaction with students. This method of instruction can be quite effective and is used frequently by on-campus tutors and writing centers because these personal meetings allow for greater attention to detail and the ability to address unique concerns of each student in a way that is not possible in larger groups. When one-on-one consultations are offered by librarians, they not only provide an opportunity to assist students, they also allow librarians to benefit from learning more about how students select and use information sources. This approach is especially helpful since we rarely have the opportunity to closely review student work in progress or the final product.

The present study is the first to explore a combination of one-on-one library re-

---

*Thomas L. Reinsfelder is a Reference/Instruction Librarian at Penn State Mont Alto; e-mail: tlr15@psu.edu. © 2012 Thomas L. Reinsfelder, Attribution-NonCommercial (<http://creativecommons.org/licenses/by-nc-sa/3.0/>) CC BY-NC*

search consultations and citation analysis, a method commonly used to evaluate the sources used by researchers. It is also one of the few attempts to use quantitative measures to evaluate the impact of individual research consultations. Earlier efforts have relied heavily on satisfaction surveys and anecdotal evidence. The impact of one-on-one student and librarian meetings was evaluated based on the quality and appropriateness of sources cited to determine if this approach to library instruction is one that is worthwhile.

Previous research using citation analysis has focused on final bibliographies to learn how students use different types of information sources. No studies were identified that examine citations or information sources while papers are still in development in order to track improvement. Further, while much of the previous work using citation analysis is concerned with the type or format of sources used, as well as dates and number of sources used, the interest in this study is primarily in assessing the quality or appropriateness of sources used, regardless of format, and measuring any improvement throughout the research process.

## Literature Review

### *Individual Research Consultations*

Numerous research consultation services using a great variety of approaches have been documented in past years. These allow students to meet with a librarian for more in-depth discussion. Some have focused on a specific population such as honors students,<sup>1</sup> graduate students,<sup>2</sup> or students in a particular course such as English.<sup>3</sup> Rowe,<sup>4</sup> Bean,<sup>5</sup> and Lee<sup>6</sup> all provide details of such an approach. Long and Shirkhande<sup>7</sup> also recognized the need for consultations with librarians as an important piece of their broader instructional strategy. Gale and Evans<sup>8</sup> suggest that individual research consultation programs are most effective when students participate voluntarily. Participation in such programs has also been observed to

be heavily dependent on the cooperation and support of faculty.<sup>9</sup>

Librarians scheduling appointments with students are better prepared to provide uninterrupted and individual attention focusing on the specific needs of just one student in ways that are not possible in a typical reference transaction.<sup>10</sup> The students also receive assistance when they have an immediate need for a solution to an information problem. The large amount of time and energy required of library staff is a common and significant challenge surrounding research consultations.<sup>11</sup> In addition, students may not show up for scheduled appointments. Further, because library research consultations usually cannot be offered to all students, it can be difficult to identify the group of students who would most benefit from such a service.<sup>12</sup>

Just as different approaches are used to provide research consultation services, different methods are used to evaluate their effectiveness. Demonstrating that it can be difficult to quantitatively prove the benefits of individual instruction, Donegan, Domas, and Deosdade<sup>13</sup> compared test scores of students who received individual instruction with those who received group instruction and observed little difference in the skills of students. Using another tactic, Bergen and MacAdam<sup>14</sup> analyzed the number and type of individuals who used the voluntary service, finding that women participated at a higher level than men, freshmen and sophomores more than juniors and seniors, and the majority of projects were longer research papers, most often in the social sciences.

While no strong quantitative data has been presented in previous literature showing a correlation between research consultations and student performance, a number of qualitative measures indicate that the practice can be valuable. User surveys and comment forms are common tools for obtaining feedback from project participants. Positive student, faculty, and librarian reactions were noted in work by

Gale and Evans,<sup>15</sup> Williamson, Blocker, and Gray,<sup>16</sup> Bean,<sup>17</sup> and Blankenship.<sup>18</sup> In another project, Rowe<sup>19</sup> concluded that, although the service was used by a small number of students, they did find it helpful. For further reading on research consultations, see the literature reviews conducted by Gale and Evans,<sup>20</sup> Bean,<sup>21</sup> and Allegri.<sup>22</sup>

### *Citation Analysis*

Librarians have used the practice of reviewing citations for many different purposes. Some of the most common are to evaluate use of library collections by better understanding the types of sources being used by researchers and to evaluate the impact of instructional efforts.

One of the most frequent reasons academic librarians study citations is to be sure local collections reflect the actual needs and interests of users. Such studies may seek to determine any noticeable trends in the level of use of electronic vs. print materials, popular vs. scholarly sources, books vs. periodicals, the dates of information sources used, as well as availability in the local collection. Citation analysis for this purpose may be done to assess the needs of a limited user population such as freshmen composition classes<sup>23</sup> or senior honors students.<sup>24</sup> In a 2003 study of first-year students, Hinchliffe et al.<sup>25</sup> noted that books and journals were being cited more than Web sites, and that popular sources were used more than scholarly items. Carlson<sup>26</sup> considered the impact of class year, course level, and academic discipline on the number and types of sources being used and found that each variable did in fact contribute to differences observed in citation behavior. Another analysis of citations in dissertations from the pre-Web and post-Web eras aimed to identify trends in various disciplines and found that, overall, the use of journals has increased while monograph usage has declined, with journals being used more frequently in science and engineering and monographs being used more frequently in the social sciences.<sup>27</sup> Others have noted

that requirements imposed by course faculty are strong predictors of citation behavior.<sup>28</sup> Knight-Davis and Sung agree that sources cited by students are "heavily influenced by the requirements regarding sources, or lack thereof, in the paper assignment."<sup>29</sup>

Instruction librarians also analyze student citations to determine the most effective teaching methods. Young and Ackerson<sup>30</sup> analyzed student papers in two groups receiving different types of library instruction and noted no difference in citation behavior in four out of five cases, while Hurst and Leonard<sup>31</sup> found that those exposed to additional library instruction used and cited more library resources. However, these differences had no impact on final grades. When Ursin, Lindsay, and Johnson<sup>32</sup> reviewed citations in student work as part of a campuswide Freshman Seminar, the low use of resources recommended by librarians indicated that instructional efforts were not as effective as previously thought. This analysis allowed librarians to consider changes to improve the program. Finally, Hovde<sup>33</sup> used citation analysis to confirm that existing library instruction practices are effective based on observations of the number of citations in student papers as well as types of sources used, especially those requiring use of library resources and search tools.

### *Measuring Quality*

"To assess the quality of research paper bibliographies, criteria and a process for rating must be formulated."<sup>34</sup> Because of this interest in measuring and improving the quality or appropriateness of sources used, it is helpful to learn how others have evaluated student citations.

Young and Ackerson<sup>35</sup> provide a thorough review of the literature-examining criteria used to evaluate the quality of student bibliographies. Some common measures include quantity of sources, format or type of source, currency, variety, relevance to the topic, authority, or simply a general rating of the level of quality or

appropriateness. As one example, Kirk<sup>36</sup> evaluated student papers using the idea of appropriateness, which took into account the reputation of a source, date, and author authority. Additionally, Kirk identified the number of references, the effective use of primary and secondary sources, and consistency in formatting of references as important items to consider. In a later assessment of instruction sessions presented by librarians and writing center personnel, Dykeman and King<sup>37</sup> analyzed the number of sources used, the amount of material by recognized authorities, the use of scholarly journals, and the variety of sources. Gratch<sup>38</sup> built on these efforts by presenting a rating scale to measure the appropriateness of sources used. Criteria included the number of sources, along with the variety, currency, and quality of information used. Here, Gratch based the idea of quality on the "reputation of the publisher, author, and any other clues that might help establish the quality of the information."<sup>39</sup> Young and Ackerson<sup>40</sup> later used the three criteria of Kohl and Wilson<sup>41</sup>—type of source, currency, and quality—to develop a rating sheet intended to help with consistency and standardization. Each citation was rated on all three criteria. Raters ranked the quality from low (popular sources) to high (scholarly sources) and assigned each item into one of five categories: conference proceedings, interviews, journals, monographs, and standards. Scores ranging from 0 (inadequate) to 3 (superior) were then assigned to each source with comparisons made among scores of students receiving different instructional methods.

The study by Young and Ackerson precedes the widespread use of electronic or Internet sources so the criteria and methods for defining high-quality or appropriate sources should be updated. When Robinson and Schlegl<sup>42</sup> studied student citations in 2004, the focus was primarily on the use and quality of electronic sources, and an electronic source classification checklist was developed.

Sources were first classified as being either high quality or low quality, based on the ability of the rater to identify the author or organization and the ability to verify legitimacy of the creator through the contact information provided. Sources found to be of high quality were then classified as being scholarly (journals and government documents) or nonscholarly (news, magazine, other).

More recently, Long and Shirkhande<sup>43</sup> created a scoring system to meet their needs in evaluating the effectiveness of library instruction. This system was described as an information literacy grading scale, with each citation scored on a 0–5 scale. Each item was evaluated for quality, variety, citation format, and information use, meaning whether or not the source was properly cited with no evidence of plagiarism. The criteria were then weighted and combined to arrive at an information literacy score for each paper. To help improve consistency in rating sources, guidelines were developed for each criterion. For example, sources were considered to be high quality if recommended by a librarian in instruction sessions or otherwise provided by the library. Sources were also considered to be high quality if they were both appropriate and authoritative for the topic being addressed.

### **Methodology and Procedures**

Two librarians, a reference/instruction librarian, and a head librarian at a branch campus of a large university evaluated the quality of sources used in drafts of undergraduate papers, met with students to make recommendations, and then reviewed the quality of sources used in the revised final product. Because the specific topic and details of an assignment must be considered when determining the appropriateness of information sources, entire papers were used rather than just the bibliography page. When student work is being evaluated in this way, the final paper's components should not be evaluated in isolation.<sup>44</sup>

The main hypothesis tested is:

Students exposed to a one-on-one consultation with a research librarian after writing a first draft will show a greater improvement in the quality of sources used on the final paper than students who did not meet with a librarian.

Library staff invited faculty to participate in a new service being offered to help improve the quality of student research papers. The introductory e-mail explained that librarians would be available to meet with students to review a draft of writing assignments. Information sources selected would be reviewed and discussed. If necessary, librarians would provide guidance on selecting additional or more appropriate information resources to help improve the final product.

Faculty were eligible to participate if they:

- had students in at least one course who would be completing a writing assignment or research paper that required the use of information sources other than assigned class readings;
- had a writing assignment or research paper that could be evaluated at two points during the project;
  1. After the first draft is completed
  2. After the paper is revised and the final draft is submitted
- were willing to share student work with librarians; and
- were willing to have students set up appointments with librarians outside class time to review the paper and information sources used.

Ten sections of various courses were identified in which students were able to participate (see table 1). IRB approval was sought and obtained with data remaining confidential both during and after the study.

In addition to observing patterns of

citation use among students who met with a librarian, a control group was established (using duplicate sections of courses in the experimental group) to observe any differences in bibliographies of students who did not meet with a librarian. These students completed the same assignments but were not asked to meet with a librarian.

While all courses required students to identify and use information resources, there was a great diversity in the types of assignments and in how faculty instructed the students. Some provided very clear guidelines for the requirements of the assignment. Others allowed more flexibility. In one course, rather than writing a complete paper, students were asked to create an annotated bibliography. In another, students worked in teams rather than individually. In yet another, students were instructed to choose from a selected list of resources and were also given guidelines on acceptable date ranges for the materials to be used. These students generally received very good scores when the quality of sources was evaluated, which is to be expected based on earlier observations of Robinson and Schlegel<sup>45</sup> and Knight-Davis and Sung.<sup>46</sup> In many cases, data were incomplete because only one paper, either the draft or final copy, was received from students. In one section, only final papers were available for this study. Additionally, a number of students assigned to the experimental group chose not to meet with a librarian. These students in the experimental group who did not meet with a librarian could have been added to a comparison group along with students in the control group. Although this would have resulted in a greater number of cases to analyze, this option was not chosen in an effort to minimize the effect of selection bias.

Faculty shared draft papers with librarians, and citations were evaluated before meeting with students based on the rating scale and guidelines developed (see figure 1). Relevance of sources was assessed based on the overall fit with the

**TABLE 1**  
**Participating Courses and Students**

|  | Course                | Assigned to control group? | # of students with draft and final paper evaluated by librarians | # of students who met with librarian | # of students who did not meet with librarian |
|--|-----------------------|----------------------------|--|--------------------------------------|---|
| 1  | English Composition   | N                          | 11   | 5                                    | 6   |
| 2  | English Composition   | N                          | 14   | 14                                   | 0   |
| 3  | Geography             | N                          | 15   | 7                                    | 8   |
| 4  | Child Development     | N                          | 17   | 17                                   | 0   |
| 5  | Occupational Therapy  | N                          | 13   | 12                                   | 1   |
| 6  | Marketing*            | N                          | 4  | 4                                    | 0   |
| 7  | Women Writers         | N                          | 2  | 2                                    | 0   |
| 8  | English Composition   | Y                          | 18   | 1                                    | 17  |
| 9  | English Composition   | Y                          | 9  | 0                                    | 9   |
| 10   | English Composition** | Y                          | 9**  | 0                                    | 9**   |
|  |                       |                            |  | 62                                   | 41  |
| * group projects                                     |                       |                            |  |                                      |   |
| ** only final papers were available for this section |                       |                            |  |                                      |   |

topic being discussed in the student's paper. Students in the experimental group were then instructed to meet with a librarian to discuss the appropriateness of information sources used. Librarians made recommendations for more appropriate or additional sources. Students in the control group were not instructed to meet with a librarian for assistance. After the paper was revised and resubmitted, the work was again evaluated by librarians. At the completion of the study, faculty were asked to answer several short questions to provide observations and perceptions of the process. Questions asked of faculty include:

- Can you describe the level and types of improvements, if any, you saw between the first and final drafts?
- How much of this improvement was related to improved use of cited sources?
- What were the students' reactions to being asked to meet with a librarian?

### Rating Scale Development

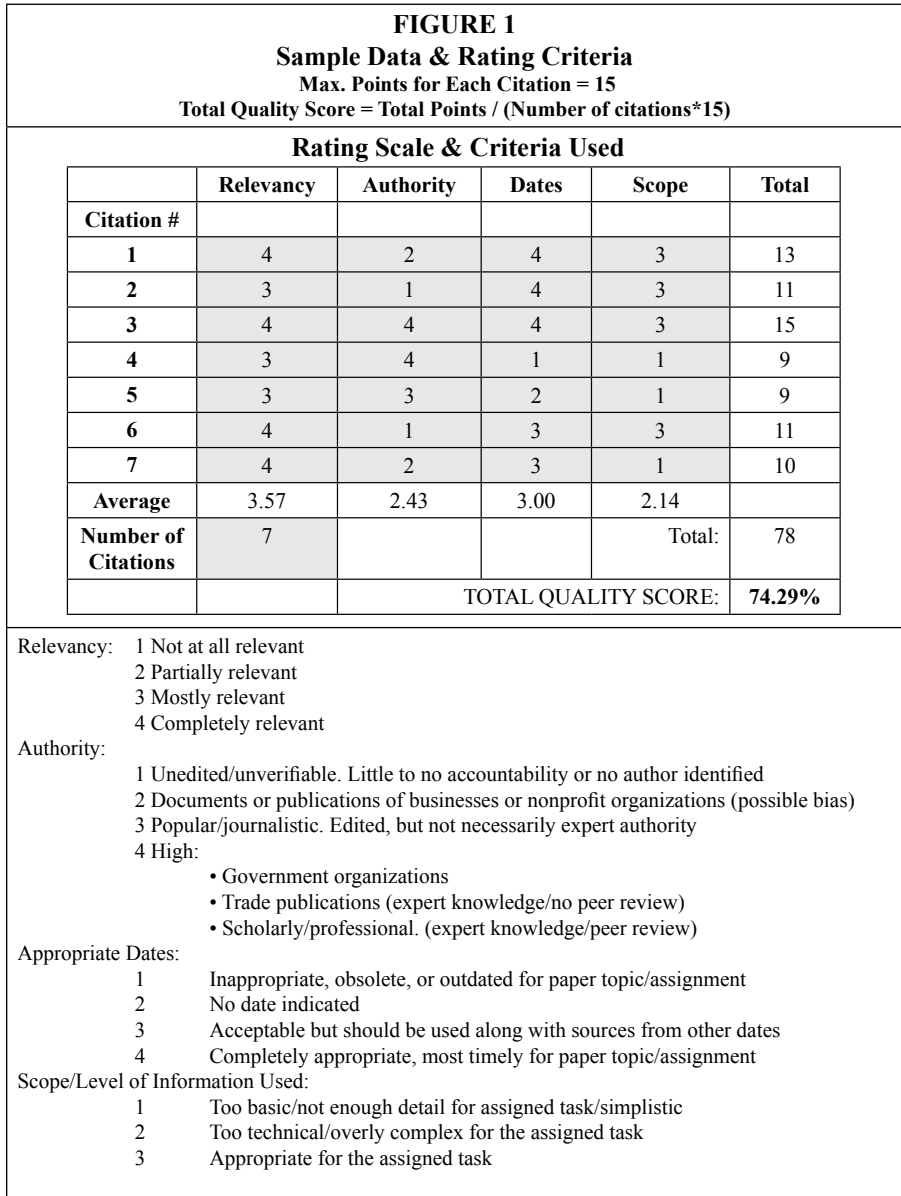
A new rating scale was developed to provide guidelines for assigning a score to each citation as well as a score for the overall bibliography (see figure 1). The instrument was constructed using elements identified in earlier works noted in the literature review, especially those concepts related to date or currency, authority, and relevance. The approach used is similar to that of Young and Ackerson,<sup>47</sup> who developed a rating sheet to evaluate sources used by students.

While the new rating scale includes some of the same ideas as measures in the past, adjustments were made to consider the appropriateness, or quality, of sources regardless of format or type of source. With so many information sources available in an electronic format, it is not always easy to separate Web sources from books, magazines, newspapers, or scholarly articles that are available electronically. Just because an item is read on a computer screen, this one characteristic is not enough to make judgments about the quality or

scholarly nature of a source. Electronic sources are often scholarly and perfectly appropriate for the work being completed. For this reason, it has been suggested that faculty “should not prohibit the use of Internet sources, but instead should implement, as stated by Davis,<sup>48</sup> ‘written and enforceable guidelines for acceptable reference sources.’”<sup>49</sup> While source type was

not addressed directly, it is accounted for to some degree when measuring authority. For example, scholarly journals and government documents are considered more authoritative than news or magazine articles and personal Web sites.

After reviewing initial papers, a measure of *scope* was added to account for the level of useful information provided by



a source. Some sources were clearly too basic for the assignment such as a brief entry from a dictionary, a Web page, or a very short encyclopedia article. On the other hand, although witnessed less frequently, some sources such as complex medical studies or legal reviews were too technical and inappropriate for the assigned task. Since the *scope* measure uses only three possible ratings, its influence on the overall score is slightly less than that of *relevancy*, *date*, and *authority*.

The rating scale provides an overall score for each paper and for each of the four criteria. The total quality score for each paper is obtained by dividing the total number of points by the total number of points possible. The maximum possible score for each citation is 15 points.

**Rating Scale Reliability**

Because two librarians were using the rating scale, and it may be used by multiple librarians in the future, it was tested for reliability among raters. Using a sample of 95 citations from one participating class, two librarians scored each citation. The assigned ratings were then evaluated for consistency. Intraclass correlations were computed using SPSS and are shown in table 2. Landis and Koch<sup>50</sup> suggest interpretations for strength of agreement among multiple raters. This analysis shows that, while the reliability among multiple raters on the *total quality score* for a paper is moderate, there was less agreement on the individual measures of *relevancy* and *scope*. On the other hand, substantial agreement was observed

for *authority*, with moderate agreement shown on *dates*.

This initial attempt to develop a rating scale is far from perfect and further testing would be beneficial to increase both validity and reliability. The most effective assessment tools and scales can be difficult and time consuming to construct, and many often mature through an extended period of testing and development. Project SAILS,<sup>51</sup> developed at Kent State University, and the READ Scale, developed by Gerlich and Berard,<sup>52</sup> are two recent examples dealing with assessment of information literacy and reference services that demonstrate how these types of efforts continue to evolve over time.

For the rating scale presented here, agreement among raters could likely be improved through more descriptive categories and criteria, along with better instructions to raters describing which ratings are most appropriate in various situations. Additionally, it would be valuable to test this instrument on a larger scale and across more diverse populations.

**Results**

Nonparametric statistical tests were chosen throughout this analysis because data did not meet the assumptions of normality required by some other common procedures. When using available data to identify differences in the quality of sources selected in draft and final papers of students who met with a librarian (experimental group), it appears that librarian input can indeed be effective. A Wilcoxon Signed Ranks Test (see table 3), the

**TABLE 2**  
**Reliability of Rating Scale among Multiple Raters**

|                         | Total Quality Score | Relevancy | Authority   | Dates    | Scope |
|-------------------------|---------------------|-----------|-------------|----------|-------|
| Intraclass Correlation* | .542                | .194      | .717        | .502     | .253  |
| Strength of Agreement** |                     |           |             |          |       |
|                         | Moderate            | Slight    | Substantial | Moderate | Fair  |

\* two-way random; absolute agreement; single measures  
 \*\*as described by Landis & Koch (1977)



**TABLE 3**  
**Draft vs. Final Papers**  
**Students Who Met with a Librarian**  
**(Experimental Group)**

| N=61                | Draft                        | Final                        | p    |
|---------------------|------------------------------|------------------------------|------|
| Total Quality Score | Mean<br>90.386<br>(SD=7.770) | Mean<br>93.676<br>(SD=6.350) | .000 |
|                     | Median 92.000                | Median 95.000                |      |
| Relevance           | Mean<br>3.770<br>(SD=.323)   | Mean<br>3.903<br>(SD=.181)   | .001 |
|                     | Median 3.830                 | Median 4.000                 |      |
| Authority           | Mean<br>3.428<br>(SD=.556)   | Mean<br>3.492<br>(SD=.555)   | .208 |
|                     | Median 3.500                 | Median 3.670                 |      |
| Dates               | Mean<br>3.681<br>(SD=.398)   | Mean<br>3.770<br>(SD=.315)   | .050 |
|                     | Median 3.860                 | Median 4.000                 |      |
| Scope               | Mean<br>2.679<br>(SD=.440)   | Mean<br>2.887<br>(SD=.251)   | .000 |
|                     | Median 3.000                 | Median 3.000                 |      |

nonparametric alternative to the paired samples t-test, showed that one-on-one consultations with a librarian during the paper writing process did result in sources of a higher quality being used on the final paper than on the draft paper. This difference was statistically significant for all measures except *authority* ( $Z=-1.260$ ,  $p=0.208$ ). Other measures were statistically significant as follows: Overall Quality ( $Z=-4.366$ ,  $p=0.000$ ), Relevance ( $Z=-3.190$ ,  $p=0.001$ ), Dates ( $Z=-1.958$ ,  $p=0.050$ ), and Scope ( $Z=-4.263$ ,  $p=0.000$ ).

When looking for differences in the quality of sources used in draft and final papers of students who did not meet with a librarian (control group), no improvement was observed in the quality of sources used. For all measures, a Wilcoxon Signed Ranks Test (see table 4) showed no statistically significant difference between

the quality of sources used on the draft paper and quality of sources used on the final paper. Results were observed as follows: Overall Quality ( $Z=-1.087$ ,  $p=0.277$ ), Relevance ( $Z=-.631$ ,  $p=0.528$ ), Authority ( $Z=-.362$ ,  $p=0.717$ ), Dates ( $Z=-.824$ ,  $p=0.410$ ), Scope ( $Z=-1.276$ ,  $p=.202$ ).

Where data was available on final papers, the quality of sources used was compared between students who met with a librarian and those who did not using a Mann-Whitney U test, the nonparametric alternative to the independent t-test. Here, a statistically significant difference was observed for *overall quality* ( $U=694.00$ ,  $p=.002$ ), *authority* ( $U=677.50$ ,  $p=.001$ ), and *dates* ( $U=848.50$ ,  $p=.039$ ), with students who met with a librarian using more appropriate sources. No difference was observed for *relevance* ( $U=1070.50$ ,  $p=.687$ ) and *scope* ( $U=933.50$ ,  $p=.092$ ).

### Testing Other Correlations

Although not the primary purpose of this study, data were further analyzed to note possible correlations among variables (see table 6). Spearman *rho* correlation coefficients were calculated for relationships between variables in both groups. For both experimental and control groups, no significant correlations exist between page length and the number of sources used on the final paper. Similarly, there is no significant correlation between the overall quality of sources and the grade received. For students assigned to the experimental group, the only significant correlation observed was for the number of sources used on the final paper and the grade received ( $r(40) = .340$ ,  $p < .05$ ). For those students assigned to the control group who did not meet with a librarian, the only significant correlation observed was for the number of pages written on the final paper and the grade received ( $r(40) = .502$ ,  $p < .01$ ).

| N=26                | Draft                        | Final                        | <i>p</i> |
|---------------------|------------------------------|------------------------------|----------|
| Total Quality Score | Mean<br>89.589<br>(SD=8.257) | Mean<br>91.215<br>(SD=5.662) | .277     |
|                     | Median 91.390                | Median 91.110                |          |
| Relevance           | Mean<br>3.895<br>(SD=.220)   | Mean<br>3.920<br>(SD=.162)   | .528     |
|                     | Median 4.000                 | Median 4.000                 |          |
| Authority           | Mean<br>3.215<br>(SD=.658)   | Mean<br>3.259<br>(SD=.500)   | .717     |
|                     | Median 3.000                 | Median 3.265                 |          |
| Dates               | Mean<br>3.625<br>(SD=.385)   | Mean<br>3.676<br>(SD=.353)   | .410     |
|                     | Median 3.670                 | Median 3.750                 |          |
| Scope               | Mean<br>2.750<br>(SD=.485)   | Mean<br>2.829<br>(SD=.308)   | .202     |
|                     | Median 3.000                 | Median 3.000                 |          |

Wilcoxon Signed Ranks Test (see table 7) indicated a statistically significant difference between the mean number of sources used on draft papers (Mean=5.41, SD=3.68) and final papers (Mean=6.39, SD=3.35) for students in the experimental group who met with a librarian. Similarly, a statistically significant difference was observed for students in the control group who did not meet with a librarian, with a greater number of sources being used on the final paper (Mean=5.19, SD=2.08) than on the draft (Mean=3.58, SD=1.27).

All students used a greater number of sources on the final paper; but, of the two groups, those meeting with a librarian used a greater mean number of sources. To determine if this number was statistically meaningful, a Mann Whitney test was used (see table 8). Here, no significant difference is indicated (U=613.00, *p*=.091).

**Number of Sources Used**

Another question that can be addressed from the data collected deals with the number of sources cited. As might be expected, students are citing more sources on final papers than on earlier drafts. A

**Feedback from Faculty**

When faculty were initially presented with the opportunity to receive assistance with the quality of student research, many were quite interested. Overall, faculty were very cooperative and encouraged their

|                     | Did Not Meet with Librarian<br>Mean Rank N=36 | Met with Librarian<br>Mean Rank N=62 | Mann Whitney U | <i>p</i> |
|---------------------|---|--------------------------------------|----------------|----------|
| Total Quality Score | 37.78   | 56.31                                | 694.00         | .002     |
| Relevance           | 48.24   | 50.23                                | 1070.50        | .687     |
| Authority           | 37.32   | 56.57                                | 677.50         | .001     |
| Dates               | 42.07   | 53.81                                | 848.50         | .039     |
| Scope               | 44.43   | 52.44                                | 933.50         | .092     |

**TABLE 6**  
**Correlations on Final Papers**

|                                    | Met with Librarian<br>(Experimental Group) |                             | Did Not Meet with Librarian<br>(Control Group) |                             |
|------------------------------------|--|-----------------------------|--|-----------------------------|
|                                    | N  | Spearman rho<br>Correlation | N  | Spearman rho<br>Correlation |
| Page Length & # of Sources         | 61   | .133                        | 36   | .300                        |
| Page Length & Grade                | 42   | .214                        | 27   | .502**                      |
| # of Sources & Grade               | 42   | .340*                       | 27   | .095                        |
| Overall Quality of Sources & Grade | 42   | .254                        | 27   | .141                        |
| *p < .05<br>**p < .01              |  |                             |  |                             |

students to work with librarians. Faculty supporting this project shared a variety of comments. Several noted that meetings with a librarian made a notable difference in the quality of sources used. Students also exhibited a better understanding of different types of sources and which are more appropriate. Nearly all agreed that this approach was worthwhile and some would like to continue to have future students meet with librarians for assistance. Others noted that this endeavor was useful in that it led to new ways of thinking about research assignments and revealed the importance of partnering with the library for assistance and expertise. Overall, little feedback was communicated to faculty from students, but some indicated that the library assistance was helpful and others seemed genuinely appreciative of the opportunity to learn about the library and new research strategies.

### Discussion

The use of a rating scale can be helpful in trying to objectively measure the quality or appropriateness of information sources used by students. Yet, there is still significant room for subjective interpretation, as it is not always clear which category a source should be assigned. For example, a book should not automatically be considered as a highly authoritative source. Many can be strongly biased and written by authors with limited professional or academic credentials. As another example, some nonprofit organizations' Web sites and publications may be strongly biased toward one point of view, while others are more professional in nature and more appropriate for scholarly work. In the case of personal Web sites, some are poorly designed and written, while others are maintained by experts with a background of impressive accomplish-

**TABLE 7**  
**Number of Sources Used**  
**Draft vs. Final**

|                                     | Mean # of<br>Sources Draft | Mean # of<br>Sources Final | Median<br>Draft | Median<br>Final | <i>p</i> |
|-------------------------------------|----------------------------|----------------------------|-----------------|-----------------|----------|
| Met with Librarian N=61             | 5.41<br>SD=3.68            | 6.39<br>SD=3.35            | 5.00            | 6.00            | .000     |
| Did Not Meet with Librarian<br>N=26 | 3.58<br>SD=1.27            | 5.19<br>SD=2.08            | 3.50            | 4.50            | .000     |

**TABLE 8**  
**Number of Sources Used on Final Papers**  
**Experimental vs. Control**

|                             | Did Not Meet w/ Librarian<br>Mean Rank<br>N=26 | Met w/ Librarian<br>Mean Rank<br>N=61 | Mann<br>Whitney U | <i>p</i> |
|-----------------------------|--|---------------------------------------|-------------------|----------|
| # of Sources<br>Final Paper | 37.08  | 46.95                                 | 613.00            | .091     |

ments. Web sites that hire freelance writers, either paid or unpaid, to contribute articles on a wide range of topics also pose special challenges. To further complicate matters, different disciplines may not agree on what constitutes a high quality source. While the humanities and social sciences may emphasize the use of books and peer-reviewed journal articles, those working in science, business, or technology fields may find it more appropriate to use sources like news reports or press releases. In all cases, it is critical to consider the authors' and publishers' background and motivations for distributing information. In meetings with students, many seemed to underestimate the importance of identifying the author and demonstrated some difficulty with this task, especially when the author was an organization rather than an individual.

By using this rating scale to evaluate the quality of sources used by students, it seems that one-on-one consultations with librarians can be an effective strategy for improving student work. Students who met with a librarian showed improvement on the measures of overall quality of sources, relevance, dates, and scope. *Authority* was the one measure for which a statistical difference was not observed. This may lead one to conclude that, for the most part, students are doing a fairly good job of identifying authoritative sources on their own, perhaps with some direction and guidance from course faculty. For students who did not meet with a librarian, there was no noticeable improvement from draft to final paper on any of the measures. When comparing the quality of sources on final papers between those

who met with a librarian and those who did not, the students meeting with a librarian scored significantly higher on *overall quality, authority, and dates*. These findings begin to provide some quantitative evidence demonstrating the positive impact of individual research consultations. However, it should be acknowledged that, although common in practice, some researchers criticize the calculation of a mean and standard deviation on data obtained from Likert type scales where initial ratings could be considered ordinal in nature, arguing that other statistical measures may be more appropriate.<sup>53</sup>

**Correlations and Number of Sources Used**

Correlations between various characteristics of student papers are interesting to consider, but the values reported here do not appear to have much practical application. However, the positive relationship between the number of sources and the grade obtained is worthy of comment. A possible explanation here is that the use of additional sources indicates a greater level of effort expended by the students, or perhaps the additional sources used by this group were recommended by a librarian and enhanced the overall quality of the final paper. But results have not been consistent across multiple studies. The 2002 study by Davis<sup>54</sup> and the 2007 work by Hurst and Leonard<sup>55</sup> saw no correlations between citation use and grades. However, Robinson and Schlegel observed that "papers with longer bibliographies tend to receive higher grades irrespective of the kinds of citations."<sup>56</sup>

When looking at another relationship, there was no indication that use of high-quality sources led to higher grades. This is likely due to the fact that student work is often graded on a number of criteria, and the quality of research resources used is only one. Davis<sup>57</sup> reached a similar conclusion, noting that assignments are also graded on the quality of writing, clarity of ideas, and the ability to address all other requirements set forth by the instructor. There is also some evidence to suggest that, even if students are able to locate high-quality information, many still struggle to successfully analyze and effectively integrate the most important ideas into their work.<sup>58</sup>

For all students, as would be expected, more sources are consistently used on final papers than on earlier drafts. It is possible that librarian interaction may lead to more sources being cited; but, in this case, no statistical evidence was found. It is also likely that the number of sources used is heavily influenced by instructions from faculty, and the requirements of the assignment as students "will meet the expectations of the professor when those expectations are clearly articulated and enforced."<sup>59</sup> Although narrowly defined guidelines for an assignment ensure the use of certain sources, it is also important for faculty and librarians to help students identify appropriate or acceptable resources without such strict parameters. Further investigation and greater collaboration with faculty is needed here to better understand the relationships between these variables. Future research may be able to make stronger claims about the nature of these relationships.

### Conclusion

For libraries considering one-on-one reference services, the rating scale presented here offers one method to measure the

impact of such a program. One of the primary drawbacks to individual instruction is that a great deal of time is required, as most meetings with students lasted from fifteen to thirty minutes. However, during this study, the time required to discuss papers and sources with students did not place an unreasonable burden on library staff. In addition, librarians were frequently able to make more meaningful connections with students by addressing the specific needs of each individual. The largest time investment came from the evaluation and scoring of citations using the rating scale and the analysis of data at the end of the semester. This type of service may be difficult to implement on a large scale across an entire campus. However, this method of instruction and assessment may be helpful in some cases, especially where librarians have a close working relationship with faculty. For those interested in efforts of a similar nature, it is suggested that librarians focus on purposefully selected groups or classes where the impact would be greatest. The existence of a coordinated campus writing program, or similar requirement, might also be used to help librarians make necessary connections with faculty.

In future research, librarians may wish to use this citation analysis approach to compare the effectiveness of research consultations to other instructional methods. Because of the potential for subjective interpretation of sources, some attention should also be paid to improving the validity and reliability of the rating scale presented here. This may be achieved through further testing and refinement of the guidelines used for assigning scores to each item. There may even be some value to creating a modified version of the rating scale to better address the needs and concerns of students and faculty in a specific subject or discipline.

---

### Notes

1. Dennis Isbell, "A Librarian Research Consultation Requirement for University Honors Students Beginning Their Theses," *College & Undergraduate Libraries* 16, no. 1 (2009): 53–57.

2. Bonnie G. Gratch and Charlene C. York, "Personalized Research Consultation Service for Graduate Students: Building a Program Based on Research Findings," *Research Strategies* 9 (Winter 1991): 4–15.
3. Jeannine Williamson, Louann Blocker, and LaVerne Gray, "The Research Assist Term Paper Consultation Program," *Tennessee Libraries* 57, no. 2 (2007), available online at [www.tnla.org/associations/5700/files/williamson.pdf](http://www.tnla.org/associations/5700/files/williamson.pdf) [accessed 9 June 2011].
4. Caroline E. Rowe, "Individual Research Consultations: A Safety Net for Patrons and Librarians," *The Southeastern Librarian* 41 (Spring 1991): 5–6.
5. Rick Bean, "Two Heads Are Better Than One: DePaul University's Research Consultation Service," in *The Seventh Off-Campus Library Services Conference Proceedings* (Mount Pleasant, Mich.: Central Michigan University, 1995), 5–15.
6. Deborah Lee, "Research Consultations: Enhancing Library Research Skills," *The Reference Librarian* 41, no. 85 (2004): 169–80.
7. Casey M. Long and Milind M. Shirkhande, "Using Citation Analysis to Evaluate and Improve Information Literacy Instruction," in *Collaborative Information Literacy Assessments: Strategies for Evaluating Teaching and Learning*, eds. Thomas P. Mackey and Trudi Jacobson (New York: Neal Schuman, 2010), 5–24.
8. Crystal D. Gale and Betty S. Evans, "Face-to-Face: The Implementation and Analysis of a Research Consultation Service," *College & Undergraduate Libraries* 14, no. 3 (2007): 85–97.
9. Lisa Blankenship, "Research Paper Counseling at UNC: Individualized Library Instruction," *Colorado Libraries* 20 (Winter 1994): 42–43.
10. Hua Yi, "Individual Research Consultation Service: An Important Part of an Information Literacy Program," *Reference Services Review* 31, no. 4 (2003): 342–50.
11. Gillian M. Debreczeny, "Coping with Numbers: Undergraduates and Individualized Term Paper Consultations," *Research Strategies* 3, no. 4 (1985): 156–63; Rowe, "Individual Research Consultations."
12. Karen A. Becker, "Individual Library Research Clinics for College Freshmen," *Research Strategies* 11, no. 4 (1993): 202–10.
13. Patricia M. Donegan, Ralph E. Domas, and John R. Deosdade, "The Comparable Effects of Term Paper Counseling and Group Instruction Sessions," *College & Research Libraries* 50 (Mar. 1989): 195–205.
14. Kathleen Bergen and Barbara MacAdam, "One-on-One: Term Paper Assistance Programs," *RQ* 24, no. 3 (1985): 333–40.
15. Gale and Evans, "Face-to-Face."
16. Williamson, Blocker, and Gray, "The Research Assist Term Paper Consultation Program."
17. Bean, "Two Heads Are Better Than One."
18. Becker, "Individual Library Research Clinics," 202–210.
19. Rowe, "Individual Research Consultations."
20. Gale and Evans, "Face-to-Face."
21. Bean, "Two Heads Are Better Than One."
22. Francesca Allegri, "One-on-One Instruction," *Medical Reference Services Quarterly* 9, no. 2 (1990): 81–84.
23. Kathleen E. Joswick, "Library Materials Use by College Freshmen: A Citation Analysis of Composition Papers," *College and Undergraduate Libraries* 1, no. 1 (1994): 43–65.
24. Leslie Kriebel and Leslie Lapham, "Transition to Electronic Resources in Undergraduate Social Science Research: A Study of Honors Theses Bibliographies, 1999–2005," *College & Research Libraries* 69, no. 3 (2008): 268–83.
25. Lisa Hinchliffe et al., "What Students Really Cite: Findings from a Content Analysis of First-Year Student Bibliographies," in *Integrating Information Literacy into the College Experience: Papers Presented at the Thirtieth LOEX Library Instruction Conference*, eds. Julia K. Nims et al. (Ann Arbor, Mich.: Pierian Press, 2003) 69–74.
26. Jake Carlson, "An Examination of Undergraduate Student Citation Behavior," *The Journal of Academic Librarianship* 32, no. 1 (2006): 14–22.
27. Thomas W. Conkling et al., "Research Material Selection in the Pre-Web and Post-Web Environments: An Interdisciplinary Study of Bibliographic Citations in Doctoral Dissertations," *The Journal of Academic Librarianship* 36, no. 1 (2010): 20–31.
28. Andrew M. Robinson and Karen Schlegl, "Student Bibliographies Improve When Professors Provide Enforceable Guidelines for Citations," *portal: Libraries and the Academy* 4, no. 2 (2004): 275–90.
29. Stacey Knight-Davis and Jan S. Sung, "Analysis of Citations in Undergraduate Papers," *College & Research Libraries* 69, no. 5 (2008): 457.
30. Virginia E. Young and Linda G. Ackerson, "Evaluations of Student Research Paper Bibliographies: Refining Evaluation Criteria," *Research Strategies* 13, no. 2 (1995): 80–93.

31. Susan Hurst and Joseph Leonard, "Garbage In, Garbage Out: The Effect of Library Instruction on the Quality of Students' Term Papers," *E-JASL: The Electronic Journal of Academic and Special Librarianship* 8, no. 1 (Spring 2007).
32. Lara Ursin, Elizabeth B. Lindsay, and Corey M. Johnson, "Assessing Library Instruction in the Freshman Seminar: A Citation Analysis Study," *Reference Services Review* 32, no. 3 (2004): 284-92.
33. Karen Hovde, "Check the Citation: Library Instruction and Student Paper Bibliographies," *Research Strategies* 17, no. 1 (1999): 3-9.
34. Bonnie Gratch, "Toward a Methodology for Evaluating Research Paper Bibliographies," *Research Strategies* 3, no. 4 (1985): 170-77.
35. Young and Ackerson, "Evaluations of Student Research Paper Bibliographies," 80-93.
36. Thomas Kirk, "Bibliographic Instruction: A Review of Research," in *Evaluating Library Use Instruction*, ed. Richard J. Beeler (Ann Arbor, Mich.: Pierian Press, 1975), 1-29.
37. Amy Dykeman and Barbara King, "Term Paper Analysis: A Proposal for Evaluating Bibliographic Instruction," *Research Strategies* 1, no. 1 (Winter 1983): 14-21.
38. Gratch, "Toward a Methodology for Evaluating Research Paper Bibliographies."
39. *Ibid.*, 173.
40. Young and Ackerson, "Evaluations of Student Research Paper Bibliographies."
41. David F. Kohl and Lizabeth A. Wilson, "Effectiveness of Course Integrated Bibliographic Instruction in Improving Coursework," *RQ* 26 (1986): 206-11.
42. Robinson and Schlegl, "Student Bibliographies Improve When Professors Provide Enforceable Guidelines for Citations," 275-290.
43. Long and Shirkhande, "Using Citation Analysis to Evaluate and Improve Information Literacy Instruction," 5-24.
44. Gratch, "Toward a Methodology for Evaluating Research Paper Bibliographies," 170-177.
45. Robinson and Schlegl, "Student Bibliographies Improve."
46. Knight-Davis and Sung, "Analysis of Citations in Undergraduate Papers," 447-58.
47. Young and Ackerson, "Evaluations of Student Research Paper Bibliographies."
48. Philip M. Davis, "Effect of the Web on Undergraduate Citation Behavior: Guiding Student Scholarship in a Networked Age," *portal: Libraries and the Academy* 3, no. 1 (2003): 41-51.
49. Robinson and Schlegl, "Student Bibliographies Improve."
50. Richard J. Landis and Gary G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics* 33, no. 1 (1977): 159-74.
51. Julia C. Blixrud, "Project SAILS: Standardized Assessment of Information Literacy Skills," *ARL Bimonthly Report On Research Library Issues & Actions*, no. 230/231 (October 2003): 18-19, available online at [www.arl.org/bm~doc/arlbr230231.pdf](http://www.arl.org/bm~doc/arlbr230231.pdf) [accessed 9 June 2011]; Kent State University, "Project SAILS (Standardized Assessment of Information Literacy Skills): History" (2011), available online at [www.projectsails.org/sails/history.php](http://www.projectsails.org/sails/history.php) [accessed 9 June 2011].
52. Bella K. Gerlich and G. Lynn Berard, "Introducing the READ Scale: Qualitative Statistics for Academic Reference Services," *Georgia Library Quarterly* 43, no. 4 (Winter 2007): 7-13; Bella K. Gerlich and G. Lynn Berard, "Testing the Viability of the READ Scale (Reference Effort Assessment Data): Qualitative Statistics for Academic Reference Services," *College & Research Libraries* 71, no. 2 (2010): 116-37.
53. Susan Jamieson, "Likert Scales: How to (Ab)use Them," *Medical Education* 38 (2004): 1217-18.
54. Philip M. Davis, "The Effect of the Web on Undergraduate Citation Behavior: A 2000 Update," *College & Research Libraries* 63, no. 1 (2002): 53-60.
55. Hurst and Leonard, "Garbage In, Garbage Out."
56. Robinson and Schlegl, "Student Bibliographies Improve."
57. Davis, "The Effect of the Web on Undergraduate Citation Behavior: A 2000 Update," 53-60.
58. Stephanie Rosenblatt, "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, no. 2 (2010): 50-61.
59. Davis, "Effect of the Web on Undergraduate Citation Behavior" (2003).

# ***Choice Reviews Online***

**[www.cro2.org](http://www.cro2.org)**

**Visit the *Choice* booth,  
#1900, at the ALA Annual  
Conference to see a  
demonstration of the  
new features on  
*Choice Reviews Online*.  
In Anaheim, CA  
June 22 - 25, 2012.**

**For a free 60-day trial to *Choice  
Reviews Online* visit  
[www.cro2.org](http://www.cro2.org).**

**For more information visit  
[www.cro2.org](http://www.cro2.org).**