World Library and Information Congress:
71th IFLA General Conference and Council

"Libraries - A voyage of discovery"

August 14th - 18th 2005, Oslo, Norway

Conference Programme: http://www.ifla.org/IV/ifla71/Programme.htm

June 9, 2005

Code Number: 078-E
Meeting: 95 SI - Statistics and Evaluation with University Libraries

Improving the quality of university libraries through citation mining and analysis using two new dissertation bibliometric assessment tools

Johanna Tuñón and **Bruce Brydges**
Nova Southeastern University
Fort Lauderdale, Florida, U.S.A.

Abstract

University libraries are becoming increasingly aware of the need to assess the quality of students' information literacy and library research skills and to use this assessment data to effectively improve the quality of university library services to graduate programs. However, libraries have had difficulties finding ways to accomplish this systematically and objectively. This study examined the relative merits of using citation analysis and evaluative bibliometric techniques to "mine" reference lists obtained from doctoral dissertations for assessment purposes. In the past, citation analysis has been used in libraries for collection development and to assess the quality of undergraduate students' library research skills. Citation analysis, however, also has the advantage of being an unobtrusive and non-invasive analytical tool that can be used to quantify students' meta-cognitive skills, beyond basic informational and procedural knowledge as captured by a pretest/posttest evaluation. This study builds on three recent works: Two studies (Beile, Boote, & Killingsworth, 2003, Haycock, 2004) used citation analysis to examine the scholarly nature of education dissertations, while another study (Green & Bowser, 2003) developed a rubric to examine the effect of a faculty/librarian collaboration on the quality of literature reviews in education

dissertations. The current study, while applying both techniques to reference lists of 144 doctoral applied dissertations' from the Child and Youth Studies program at Nova Southeastern University, goes the next step by creating a method of citation analysis for the purpose of gathering evaluative, bibliometric data. The writers developed an objective rubric that awarded points for currency, type of document, and certain document-specific criteria, while the second rubric employed a subjective assessment based on five criteria (number and variety of types of documents cited, depth of understanding as demonstrated through the inclusion of theoretical and background documents as well as scholarliness, currency, and relevancy of the resources). Qualitative descriptors were used to score the criteria on an eight-point scale. A comparison of the two overall scores provides evaluative evidence of the quality of students' library research skills as demonstrated in this graduate capstone endeavor. The data was also analyzed to generalize knowledge, predict the quality and structure of dissertation literature reviewing, and to improve university library services and resources.

Background

University libraries are becoming increasingly aware of the need to assess the quality of students' information literacy and library research skills and to use this assessment data to effectively improve the quality of university library services to graduate programs. However, libraries have had difficulties finding systematic and objective ways to accomplish this. Libraries consider themselves as part of the academic process of promoting scholarship and learning in students, but the data collected by libraries has not always been able to document a correlation between these two factors (Dugan & Hernon, 2002, p. 376). Part of the problem is that there has been a disconnect between the statistics being gathered and proof that libraries actually contributed to students acquiring higher-level library research skills. Moreover, measuring the learning outcomes for library research skills has been a challenge because these types of skills are considered institutional rather than simply library objectives (Bertot & McClure, 2003) and therefore a "meta-outcome in the learning process" for the institution as a whole (Ratteray, 2002, p. 370).

Many academic libraries (Zuniga, Webcast discussion, 2004) have begun looking at student projects, theses, and dissertations as a source of evidence of library research skills. Citation analysis has been used for decades as an analytical tool for examining reference citations in bibliometric studies. Historically, "citation analysis" has been defined as the study of citations to and from documents within a literature (*International Encyclopedia of Information and Library Science,* 2002*)* while "bibliometrics*"* is broader in scope and is defined as the "study and measurement of the publication patterns of all forms of written communication and their authors" (Potter, 1981). Citations are particularly appealing because they can be treated as objects that can be described and counted. Moreover, because dissertation and theses reference lists can be obtained without the participation of the authors, the data is relatively easy to collect and has the benefit of not being uncontaminated by participant responses and opinions. However, in spite of the fact that citations can be used as a partial indicator of quality of students' references, most citation analyses of dissertations have been bibliometric studies conducted to help libraries to make informed collection development decisions (Edwards, 1999; Gooden, 2001; Herubel, 1991; Kuyper-Rushing, 1999). Only a few dissertations have focused on the social sciences in general (Hovde, 2000; 1990; Sylvia, 1998; Thomas, 2000) or education in

particular (Budd, 1988; Okiy, 1991; Cole, 1992; Glynn, 1995; Iya, 1996; Beile, Boote, & Killingsworth, 2003; Haycock, 2004).

Some studies have taken citation analysis to the next level to help examine the quality of references lists, usually in conjunction with subjective rubrics. These studies were able to quantify students' actual ability to manipulate library tools rather than simply surveying student attitudes (Hovde, 2000). A number of studies have used citation analysis of undergraduate students' bibliographies (Dykeman & King, 1983; Hurst & Leonard, 2005; Kohl & Wilson, 1986; Magrill & St. Clair, 1990; Ackerson, Howard, & Young, 1991; Young & Ackerson, 1995; Oppenheim & Smith, 2001) as an indicator of the success of library training. Malone and Videon (1997) and Hovde used citation analysis to do quantitative analyses of works cited. Although Green and Bowser (2003) developed a good rubric for master's education thesis that was not in conjunction with citation analysis, only Beile, Boote, and Killingsworth (2003) used citation analysis in conjunction with a subjective rubric to assess education dissertations.

When doing a citation analysis, dissertations have proven to be particularly appealing to use for assessing quality of reference lists because they are the "culminating experience of doctoral training [that] is crystallized in a dissertation" (Herubel, 1991, p. 65). As Buttlar (1999) noted, "the doctoral dissertation is evidence of the author's ability to engage in an extensive scholarly endeavor" (p. 228). Moreover, dissertations are important because they provide libraries with evidence of meta-cognitive skills that go beyond information and procedural knowledge acquired in pretest/posttests.

At Nova Southeastern University in Fort Lauderdale, Florida, a faculty member in the Fischler School of Education and Human Services and a librarian in the Alvin Sherman Library, Research, and Information Technology Center decided to collaborate to look at the quality of dissertation reference lists results. The Sherman Library was preparing for reaffirmation of accreditation in 2007 and was interested in looking at dissertation reference lists for evidence of library research skills. The Fischler School of Education and Human Services was interested in assessing the quality of their dissertations. This effort came at an opportune time. The issue of the quality of Ed.D. dissertations had suddenly became a "hot" topic in the United States in 2005 with the release of a report questioning the quality of Ed.D. programs (Jacobson, 2005; Levine, 2005). As a result, Nova Southeastern University (NSU) located in Fort Lauderdale, Florida, as one of institutions of higher education that graduated the highest numbers of Ed.D. students in the United States, became even more interested in the development of valid and reliable measures to quantify quality dissertation reference lists.

Research Questions

This study looked at the following questions:

- Can bibliometric/citation analysis of reference lists be conducted reliably using either objective or subjective criteria or a combination of the two?
- Can the objective and subjective criteria developed by the researchers be used to assess the quality of a graduate dissertation/thesis reference list adequately?
- Can acceptable inter-rater reliability between subject expert and library practitioner be established?
- Does different methods of library training impact the quality of the dissertation reference lists?
- Are there differences in the quality of dissertation reference lists produced by students attending classes locally and at field-based sites?

- To what degree can citations be mined for additional evidence of students' library research skills using electronic resources?

Methodology and Data Analysis

The researchers developed objective and subjective criteria for citation analysis. They started by conducting a search of the literature to identify objective and subjective criteria used in past bibliometric/citation analyses. The authors were interested in using citation analysis to go beyond simple descriptive bibliometrics to using an evaluative bibliometric application to gather quantitative measurements and statistical data on the condition or character of the applied dissertations of CYS students. The researchers developed a subjective rubric using elements from Kohl and Wilson (1986), Beile, Boote, and Killingsworth (2003), and Green and Bower's (2003) rubrics. They used a four-point scale: 1 for inadequate, 2 for marginally adequate, 3 for adequate, and 4 for superior, but they awarded half points on occasion. A faculty-librarian team of co-raters was utilized (Dykeman & King, 1983; Kohl & Wilson, 1986; Young & Ackerson, 1995; Malone & Videon, 1997; Beile, Boote, and Killingsworth, 2003). No effort was made, however, to look at the way the resources were used in the dissertation as a whole or more specifically in the literature review.

Reference lists from applied dissertations produced by students in the Child and Youth Studies (CYS) program were used in the study for a number of reasons. First of all, the strong history of cooperation between the library and the Child and Youth Studies Doctoral Program made using the reference lists from the culminating piece of doctoral research for graduate students in CYS an ideal place to unobtrusively observe, assess, and document students' acquired research skills, particularly since the program had traditionally placed a high emphasis on the review of the research literature because its interpretation of the nature of applied research. Secondly, there had been a unique history of collaboration between the Sherman Library and CYS personnel to provide library training that has taken a number of forms over the last few years.

The five criteria used in the subjective rubric included:

- breadth of resources that considered the number of citations and the variety of citations cited were considered,
- depth of understanding as demonstrated through the citing of historical and theoretical background resources,
- depth of scholarliness based on the use of primary resources and peer-reviewed resources, empirical research, and seminal or landmark studies,
- currency,
- and relevancy of the resources cited for the topic being researched.

When developing the objective rubric, the researchers had less success in identifying relevant studies. A review of the literature identified a number of studies that used citation analyses to count the frequency that various types of resources were cited. Some only counted journal citations (Chambers & Healey, 1973; Thomas, 2000) while others did count books and a variety of other types of resources (Glenn, 1995; Gooden, 2001; Haycock, 2004; Hovde, 2000; Kuyper- Rushing, 1999; Malone & Videon, 1997). However, none of the studies went beyond simply counting the frequency of various types of resources.

Because the researchers wanted to be able to quantify "student awareness of specialized [types] of resources" (Hovde, 2000, p. 5), they adopted elements from several classification schemes (Radhakrishna, 1994; Glynn, 1995; Buttlar, 1999) for a typology of resources. They sorted citations into the following categories: (1) journal articles, (2) books and book chapters from commercial publishers and university presses, (3) conference papers/proceedings, (4) reports and other gray literature by government agencies, universities, associations, and foundations, (5) dissertations, theses, practicums, and action-based research projects, (6) newspapers, (7) ERIC documents that did not fall into previous categories, (8) laws and court cases, and (9) miscellaneous other documents as a "catch all" for everything from tests, unpublished manuscripts, videos, CDs, and computer software to raw data, poster sessions, PowerPoint presentations, brochures, and more. The researchers then developed a weighting scale using objective criteria that included points for currency, scholarly types of publications, and journal characteristics including peer-reviewed, type of periodical, and whether the publication was rated as academic/scholarly. The criteria used were enhanced by professional input and validated by collection development, reference, and instructional librarians in the Sherman Library. The resulting rubric assigned points based on a completely objective set of criteria. See Table 1 for details.

Table 1

	points	**≥ 3**	**≥ 10**	**Max pts**
Dissertations * (published and unpublished)	2	.3	.2	**2.5**
Theses/practicums/action-based research*	1	.3	.2	**1.5**
Periodicals (magazines, trade journals)*	0	.3	.2	**.5**
Scholarly periodicals*	1.5 (max)	.3	.2	**2.0**
Journals +.3				
Academic/scholarly +.2				
Peer-reviewed +1				
Books/book chapters (not scholarly)	0	.3	.2	**.5**
Books/book chap. - scholarly publishers	1	.3	.2	**1.5**
Books/book chap. - academic presses	1	.3	.3	**1.5**
Reports (gov. agencies, foundations, associations, universities, etc.)*	1	.3	.2	**1.5**
Conference papers and proceedings* (published and unpublished)	1	.3	.2	**1.5**
Government laws/legal cases	1	0	0	**1.0**
ERIC ED documents*	.5	0	0	**.5**
Newspapers*	0	0	0	**0**
Web sites*	0	0	0	**0**
Miscellaneous*	0	0	0	**0**

***** Documents that fit two or more categories were included in category with higher weight

Working with citations gathered from references lists rather than a database required a great deal of work. Reference citations were scanned, and the citations normalized and sorted in Excel. The researchers then identified and categorized types of citations in dissertation reference lists. When resources fitted into more than one category, the researchers decided count function rather than form when possible. Thus, a government report that was retrieved full-text online from the ERIC database was counted as a report rather than a Web page or ERIC document. The objective rubric was used to generate an algorithm in Access. A cover page with basic data about number and types of citations and their currency was generated. Two co-raters, a program professor in CYS and a librarian, used the cover sheet when scored the subjective rubrics. (See

Appendix B for sample cover sheet.) The data was also used to produce the objective scores but only after all reference lists had been scored using the subjective rubric. (See Appendix C for sample of a page with objective scores.) A total of 10,029 citations in 144 dissertation reference lists were digitalized, normalized, sorted, and processed in Access to produce the cover sheets with bibliometric data and the objective rubric scores.

Once the two rubrics were scored, the researchers analyzed a number of things. Inter-rater reliability was assessed, the correlation was quite high, due in part to the fact that the raters worked hard to establish consensus on the criteria used. They had initially practiced on several reference lists to come to agreement on the meaning of qualitative indicators such as "disproportionate, limited, reasonable, exhaustive, etc." Inter-rater reliability was tested using a Spearman, two-way mixed effects model of the intraclass nonparametric correlation coefficient in SPSS version 13.0. A correlation coefficient of .978 was found to be significant at the 0.01 level (2 tailed). In contrast, Kohl and Wilson (1986) had used a Pearson Correlation two-tailed test for ratings of bibliographies by the librarian and instructor and found a coefficient of .679 at the significance level of .001.

Strong and statistically significant correlations existed between the library specialist's and faculty subject specialist's assessments and the objective rubric total. The agreement between the objective total and subjective reviewer's totals are much more spread out. (Table 2)

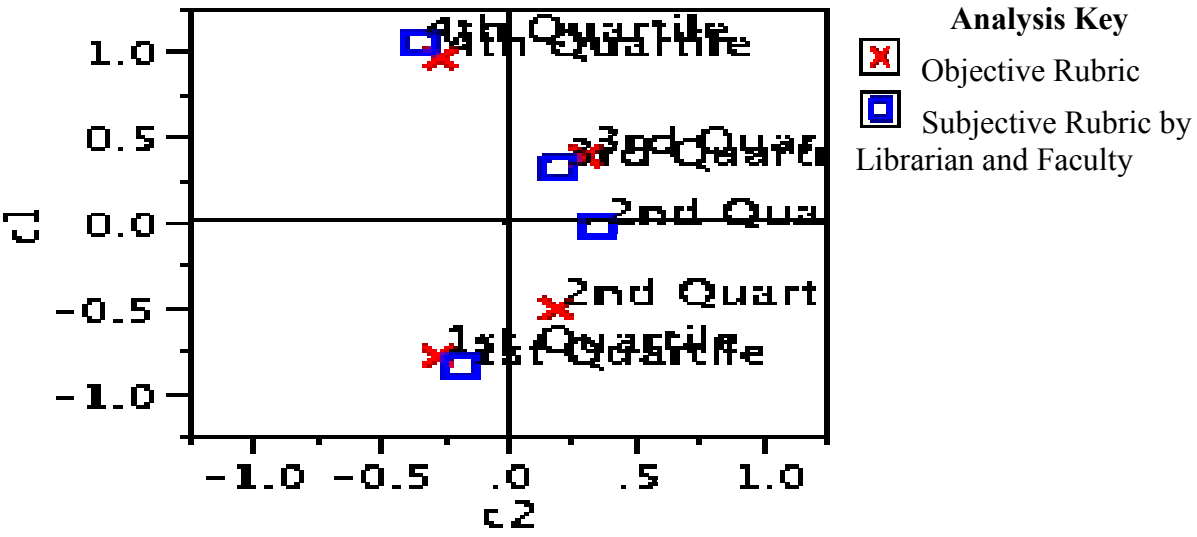
Table 2 Univariate Simple Statistics

Column	N	Mean	Std Dev	Sum	Minimum	Maximum
Librarian	143	14.6748	2.0847	2098.50	9.0000	19.000
Faculty	143	14.5839	2.1963	2085.50	8.5000	19.000
Objective	143	73.5944	27.4539	10524.0	20.1000	157.60

A correspondence analysis by quartiles revealed that the agreement between the librarian and the objective rubric assessment was fair. A Kappa statistic and the agreement between the objective rubric and the librarian's quartiles was 29% which is considered fair. On the correspondence plot, the greatest disagreement occurred within the 2nd quartile. It appears that the objective measure placed persons in the second quartile more closely with those in the 1st quartile, while the librarian placed them more closely with people in the 3rd quartile.

Using quartiles, Table 3 showed that the agreement between the faculty subject specialist and the objective rubric was a little better than the librarian's and the objective rubric agreement. A Kappa statistic and the agreement between the objective rubric and the faculty subject specialist quartiles was 33%, and that is also considered fair. On the correspondence plot, the greatest disagreement again occurred within the 2nd quartile. The objective measure placed persons in the second quartile more closely with those in the 1st quartile, while the faculty subject specialist placed them more closely with people in the 3rd quartile. So, overall the grading rubrics closely

Table 3 Correspondence Analysis



paralleled the objective measure. A bit more fine-tuning with the second quartile would improve the statistical agreement although perhaps at the risk of losing the value of subjective assessment. Both reviewers were better able to score a 1st quartile, but they experienced more difficulty differentiating between second and third quartile measures. Clearly, people in the fourth quartile are easy to spot. This is fairly typical in most subjective assessment methods.

Table 4 provided a graphic representation of the symmetry between the two methods of scoring. Because the sample was large, it also established a contingency table that demonstrated the statistical significance of the relationship and the reliability of these tools when assessing the quality of dissertation reference lists.

Table 4 Mosaic Plot of Contingency Table Quartiles of Objective Rubric by Quartiles of Faculty

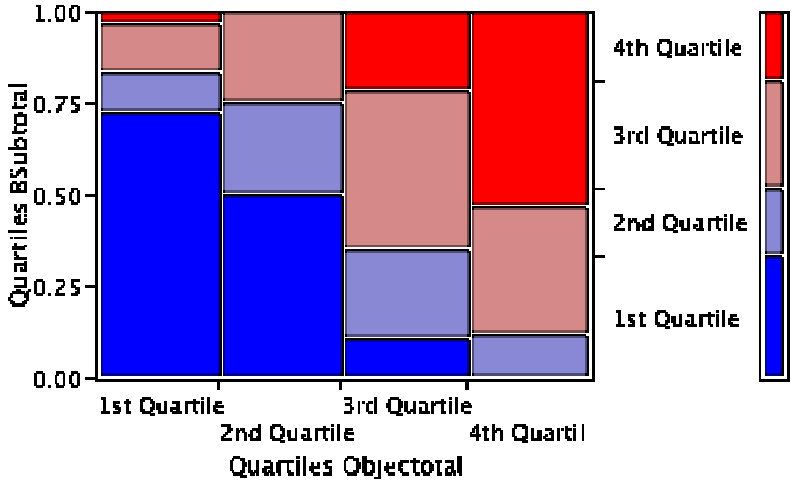


Table 5 Scatter Plot Matrix

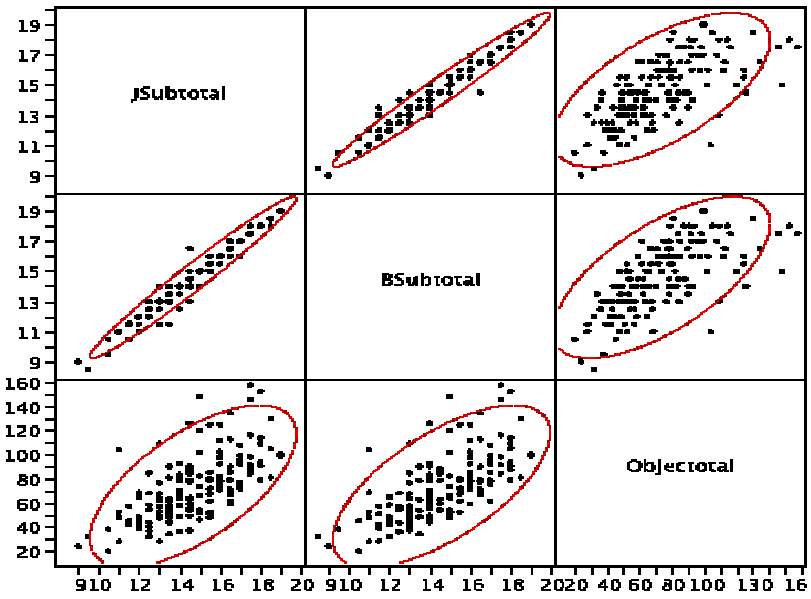


Table 5 shown above documented a positive correlation in both methods of scoring although the agreement between the objective total and subjective reviewer's totals were much more spread out. This could be attributed to the result of the objective rubric's strong relationship to quantity as discussed above and demonstrated in the resulting normal curve (Table 6)

The dissertation citations were scored for breadth, depth of understanding, depth of scholarliness, currency, relevancy and an overall quality score. Across all coded citations, the mean statistic for breadth was 2.57 (SD = .95), skewness was -.150 (SE = .202), and kurtosis was -.878 (SE = .404); the mean statistic for depth of understanding was 2.68 (SD = .81), skewness was -.326, and kurtosis was -1.069; depth of scholarliness was 3.03 (SD = .75), skewness was -.070, and kurtosis was -1.29; currency was 2.59 (SD = .67), skewness was -.209, and kurtosis was -.150; relevancy was 3.72 (SD = .50), skewness was -1.07, and kurtosis was 2.19. The overall quality statistics were m = 14.6 (SD = 2.197), skewness was -.206, and kurtosis was -.38. (See Table 7)

Table 6 Normal Curve Quality Scoring Distribution for Objective Rubric

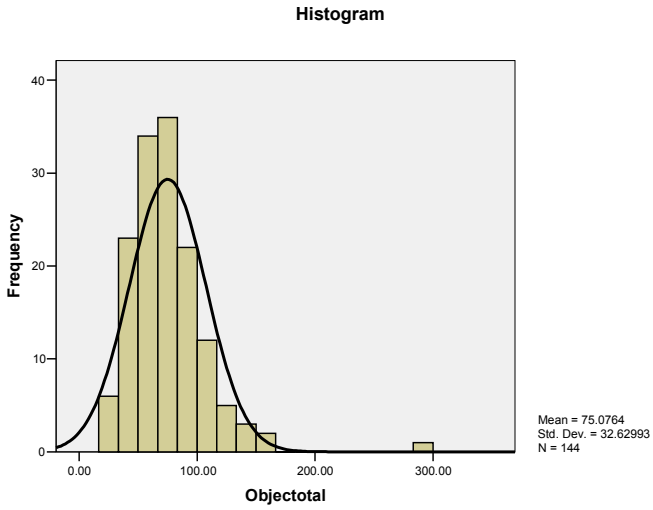
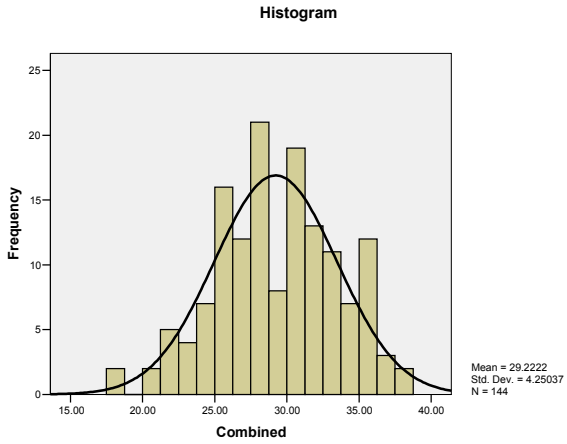


Table 7 Normal Curve Quality Scoring Distribution for Subjective Rubric



A Profile of CYS Dissertation Reference Lists

A total of 10,029 citations from the 144 dissertation reference lists were analyzed. Since all dissertations available had been included in the sample, there was no sampling error. As Table 8 demonstrates, the resources cited in the dissertation reference lists embodied a fairly diverse range of material types, but they were still heavily skewed toward only a few categories. A total of 69% of all resources cited by CYS students were from periodicals while only 18% were from books, and the rest from a variety of sources. The distribution of journals and books used by CYS students were significantly different from the findings in the studies by Beile, Boote, and Killingsworth (2003) and Haycock (2004). Beile et al. reported journal articles used 45% of the time, monographs 33% of the time, and other resources 18.3% of the time. Haycock reported 44% journals and 56% monographs and reports. The high percentage of periodicals used by CYS students in this study raised questions about reason for the differences in usage

patterns between NSU's doctoral education students and Ed.D. students at the other institutions. The co-investigators felt that further research was needed to establish whether this was specific to CYS students or was typical for all Ed.D. students at NSU.

Table 8 Citations by Document Type

Periodical articles*	6902	69%
Dissertations/theses/practicums*	76	1%
Books	1800	18%
Conference proceedings*	112	1%
Web sites*	151	2%
ERIC ED documents*	164	2%
Newspaper articles*	89	1%
Legal documents and laws	28	1/4%
Reports*	496	5%
Miscellaneous	211	2%
Total	10029	

* Documents that fit two or more categories were included in category with higher weight

Table 8 provided a profile of the average Child and Youth Studies dissertation reference list. Of the 69.6 references, 47.7 citations or 68.7% on average were from periodicals (journals, magazines, trade publications, newspapers, etc.). Of the periodicals cited, an average of 33.2 were from peer-reviewed publications, and 23 were considered scholarly as defined in *Ulrich's Directory of Periodicals'* classification scheme. On average, the typical CYS dissertation cited 12.7 books or chapters from books, 3.4 government documents, 0.8 conference papers, 0.5 dissertations and theses, 1 Web site, 0.2 laws, 0.6 newspapers and magazines, and 0.2 ERIC documents. (It should be noted that citations were only counted as ERIC documents when they did not fit in other categories such as conference papers or government reports.) As for currency, an average of 24.3% of all of the citations were from resources published within 3 years from the date of the dissertation's completion, 54.1% were from documents published between 3 and 10 years from completion date of their dissertation, and 20.9% were from documents published 11 or more years from the completion date of the dissertation. Only 53 students out of 144 cited one or more conference papers, and 23 students cited one or more dissertations. The total number of citation per dissertations ranged from a low of 23 to a high of more than 250. (M = 69.6, SD = 29.2)

Methods of Library Research Skill Training:

The co-investigators analyzed CYS cohorts based on the types of library instruction received. Students in Clusters 93 to 98 (n = 69) received a three-step training process (Tunon, 1999) while students in Clusters 99 to 109/113 (n = 74) received the traditional "one-shot" library training session early in the program that lasted 1 ½ hours. As a result, the reference lists were divided by type of intervention into Cohorts 1 and 2. The researchers compiled the reference lists of the two intervention groups (three-step and one-shot training) as well as the dates when the dissertations were completed and approved. (Table 9)

Analysis of data for the differing methods of library training was conducted using a Levene's test for equality of variance and a t-test for equality of means. Although the mean score of the second cohort (the clusters who were given the one-shot method of training) was slightly higher than for the first cohort (the clusters who received a more in-depth three-part training), the two-tailed score of .246 indicated that there was no statistically significant difference between the two methods of library training and the scores received by students in either of the training cohorts (Table 10). This was attributed to the fact that the three-part training, although more extensive than the one-shot training, was not integrated into the course work as originally planned. In addition, both groups of students were impacted by improvements taking place to the Web during that time period and access to what Rogers (2001) termed a "critical mass" of electronic journal by the late 1990s. To look into this further, the co-investigators plan to look at the results of a third cohort that received training in a two-day, team-taught workshop in future research to see if the method of training produces improved quality of citations or not.

Table 9 Means and Standard Deviations

Level	Number	Mean	Std Dev	Std Err Mean	Lower 95%	Upper 95%
Cohort One	69	71.9638	26.0657	3.1379	65.702	78.225
Cohort Two	74	75.1149	28.7819	3.3458	68.447	81.783

Table 10 Assuming Unequal Variances

Difference	-3.151	t Ratio	-0.68695
Std Err Dif	4.587	DF	140.8847
Upper CL Dif	5.917	**Prob > |t|**	**0.4932**
Lower CL Dif	-12.219	Prob > t	0.7534
Confidence	0.95	Prob < t	0.**2466**

A Comparison of Distance and Local Cohorts:

The researchers wanted to find out if there were any differences in the quality of reference lists between students who attended classes locally and had easy access to the print collection at the Sherman Library and students that attended classes at field-based sites at locations outside the three-county area of south Florida. To assess this, they analyzed results from the 144 dissertation reference lists by dividing the students into a local (N=51) cohort of students attending classes in the three-county area of south Florida and a field-based (N=93) cohort who students who attended classes at sites outside the three-county area. As Table 11 demonstrated, the researchers found no significant difference between students in local and field-based sites. What was notable was that the co-investigators found almost equal probability across quartiles for each group. Likewise, assuming unequal variances, no statistically significant differences were found between the two groups. (See Table 12.)

Table 11 Ordinal Logistic Regression

Probability Reported Quartile for Subjective scoring

Group	Probability of First Quartile	Probability of Second Quartile	Probability of Third Quartile	Probability of Fourth Quartile
Distance Students	24.1%	24.8%	26.2%	24.9%
Local Students	26.3%	25.5%	25.5%	22.7%

Table 12 Ordinal Logistic Fit for Quartiles Subjective Scoring Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
Group	1	1	0.16	0.68

Use of Bibliometric Data for Citation "Mining" and the Limitations of This Method

Since the citations had been digitalized and sorted already, the bibliometric data from journal citations were "mined" for information. Like a number of previous studies using citation analysis (Budd, 1988; Okiy, 1991; Cole, 1992; Glynn, 1995; Iya, 1996; Beile, Boote, & Killingsworth, 2003; Haycock, 2004), the most frequently cited titles were identified. Because Child and Youth Studies was an interdisciplinary field, the co-investigators also looked to see what percentage of the most frequently used journals were categorized in the area of child and youth and/or education. All but one of the 18 journals most frequently cited in Table 13 were classified as education resources while only 3 journals were considered child and youth journals, and 1 title, *School Psychology Review*, was not included in either category in spite of the title. The researchers analyzed the journals using several other characteristics identified in *Ulrich's Directory of Periodicals*. Only 13 of the top 18 journals were categorized as journals, and 14 were identified as being peer reviewed.

Table 13 Ulrich's Characteristics of Most Frequently Cited Periodicals

	Journal/ Magazine	Academic/ Scholarly	Peer-Reviewed	Education Child/Youth
Educational Leadership	j	a/s	-	ed
Phi Delta Kappan	j	a/s	-	ed
Exceptional Children	mag	a/s	pr	ed/cy
J. of Research on Comp. in Educ.	-	a/s	pr	ed
Reading Teacher	j/mag	a/s	pr	ed
Intervention in School & Clinic	j	a/s	pr	ed
Clearing House	j	a/s	pr	ed
Journal of Learning Disabilities	j	a/s	pr	ed
Remedial and Special Education	j/mag	a/s	pr	ed
Journal of Educational Research	j	a/s	pr	ed
Adolescence	j	a/s	pr	ed/cy
Education	mag	-	pr	ed
Young Children	j	a/s	pr	ed/cy
Preventing School Failure	j	a/s	pr	ed
Educational Digest	mag	a/s	-	ed
NASSP Bulletin	j	a/s	-	ed
Journal of Educational Psychology	j	a/s	pr	ed
School Psychology Review	j	a/s	pr	psy

As with data mining, the structure and format of content really mattered when the researchers used citation analysis for more than simply tools for assessing a journal collection. Because the migration of resources from print to online in the 1990s, the co-investigators were interested in using the citations to unobtrusively examine students' patterns of use of scholarly and popular resources retrieved online. Although, as noted previously, resources had been sorted by function rather than format, bibliometric data had been noted about online. The citations were mined to identify the range of resources in these citations from scholarly resources (e-books, journal articles, reports from government agencies, foundations, associations, and institutions of higher education, ERIC documents, dissertations and theses, and conference papers) to unscholarly resources (newspapers and unpublished documents posted on Web sites) constitute a growing proportion of materials.

Table 14 Where Resources Were Retrieved

	Database	URLs	ED #s
Journal articles	367	54	0
Reports	44	56	106
Web sites	0	190	0
Newspaper articles	6	26	0
Dissertations/practicums/theses	6	1	10
Conference papers	0	4	18
ERIC documents	4	12	2
Laws/legal cases	0	1	0
Books	0	0	0
Totals	**427**	**344**	**136**

Although the citations documented that a variety of types of scholarly resources were retrieved online including journal articles, dissertations, conference papers, reports, and ERIC documents, less than 1% of all citations of the 10,029 citations were formatted as having been retrieved online. When online resources were used, CYS students were most likely to use them for government census statistics, unpublished university documents, and reports from various types of agencies. As Table 14 demonstrates, a number of document types had ERIC ED document numbers, but even though the majority of these were available full text online, they rarely included retrieval statements however, less than half of one percent of the journal articles included retrieval statement. Of the 339 unique journal titles cited as being obtained online, only 4 were obtained from open access journals. The rest were from subscription databases or journal Web sites.

Some red flags were raised, however, about the reliability of the low percentages of resources cited as being retrieved online. For example, the fact that only 367 journal articles out of 6,902 (less than half of one percent of all articles) were cited as having been retrieved from NSU's subscription vendors, ProQuest, Wilson, and Gale/InfoTrac raised questions about students' knowledge of or attention to the American Psychological Association's citation formatting guidelines for resources retrieved online. Only 36 out 144 students used any retrieval statements of any kind. Of these, 25 cited articles retrieved from Web sites, and 22 included retrieval statements for articles retrieved from one or more full-text databases. The fact that all 144 CYS students had been introduced to the Wilson Web Education Full Text database during library instruction sessions made the reliability of these numbers particularly suspect. However, these findings were in line with several previous studies (Beile, Boote, & Killingsworth, 2003; Gooden, 2001; Malone & Videon, 1997) that also questioned students' lack of retrieval statements.

In order to test the hypothesis that the low number of citations might have been due to lack of knowledge about formatting conventions, the researchers checked the reference lists of 19 doctoral students in another Ed.D. program at NSU who had received more intense training on searching subscription databases and APA formatting protocols. The researchers were disappointed to find that only 4 out of the 19 students actually used retrieval statements in their dissertation reference lists.

There are several possible explanations for the lack of included retrieval statements by NSU students in the two Ed.D. programs. Simply sloppiness of the part of students in paying attention to the nuances of APA formatting rules was one possible explanation, but it was also possible that the fact that APA rules for citing resources retrieved online had been in flux during the previous decade may have resulted in confusion on the part of students. Still a third possibility was that dissertation committee members were not always up about how to cite resources retrieved online and may have advised students not to include retrieval statements in their citations.

Whatever the source of the problem, the lack of wide-spread inclusion of retrieval statements was disappointing for the co-investigators. Retrieval statements would have served as a "performance measure" (Mercer, 2000, para. 2) of students' knowledge of how to use the library's online resources that could have used to empirical quantify students' use of databases for both library and accreditation assessment purposes. The lack of consistent use of retrieval statements in citations, coupled with errors caused inaccurate bibliometric information in the citations, points to one major limitation of using citation analysis for bibliometric data.

Conclusions

Citation analysis used in conjunction with rubrics can be an effective technique for assessing the quality of reference lists. The researchers were able to develop and validate two very different but reliable tools that can be used with dissertation reference lists in education as well as other subject areas. The mechanical but objective nature of the one rubric balances the nuanced results of the subjective rubric. Past studies have already documented how lists of frequently cited journals and books can be used to access a library's collection. However, as print resources migrate online, it is less clear what insights can be mined from citations about the ways resources emerging from networked digital environments may be used for research and learning. As long as students do not include retrieval statements, the bibliometric information about students' use of electronic resources will be murky at best. Nevertheless, the two assessment tools developed and validated in this study provide libraries and academic institutions with two effective tools for assessing the quality doctoral students' higher-level library research skills and inform the efforts of academic programs and libraries alike.

References

Ackerson, L. G., Howard, J. G., & Young, V. E. (1991). Assessing the relationship between library instruction methods and the quality of undergraduate research. *Research Strategies,* 9, 139-141.

Beile, P. M., Boote, D. N., & Killingsworth, E. K. (2003, April). *Characteristics of educational doctoral dissertation references: An inter-institutional analysis of review of literature citations.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED478598) Retrieved November 20, 2004, from ERIC database.

Bertot, J. C., & McClure, C. R. (2003). Outcomes assessment in the networked environment: Research questions, issues, considerations, and moving forward. *Library Trends, 51*(4), 590-613.

Budd, J. (1988). A bibliometric analysis of higher education literature. Research in Higher Education, 28, 180-190.

Buttlar, L. (1999). Information sources in library and information science doctoral research. *Library & Information Science Research, 21*(2). 227-245.

Jacobson, J. (2005, March 25). Reports call for abolition of Ed.D. degree and overhaul of education schools. *Chronicle of Higher Education, 52*(29), A24. Retrieved March 29, 2005, from http://chronicle.com/temp/reprint.php?id=u9s4moadqbi8oadqzqmibwiwn343ry

Chambers, G. R., & Healey, J. S. (1973). Journal citations in master's theses: One measurement of a journal collection. *Journal of the American Society for Information Science, 24*, 397-401.

Cole, K. S. (1992). Doctoral students in education and selecting resources for the literature review. *Dissertation Abstracts International, 53*(07), 2142A. (UMI No. AAT 9235619). Retrieved April 13, 2003, from Digital Dissertations database.

Dugan, R. E., & Hernon, P. (2002). Outcomes assessment: Not synonymous with inputs and outputs. *The Journal of Academic Librarianship, 28,* 376-380.

Dykeman, A., & King, B. (1983). Term paper analysis: A proposal for evaluating bibliographic instruction. *Research Strategies, 1*(1), 14-21.

Levine, A. (2005). *Educating school leaders*. Report 1 from the Education Schools Project. Retrieved April 17, 2005, from http://www.edschools.org/pdf/Final313.pdf

Edwards, S. (1994). Bibliographic Instruction Research: An analysis of the journal literature from 1977 to 1991. *Research Strategies, 12*(2), 68-78.

Glenn, D. L. (1995). A citation analysis of master's and education specialist theses and research papers by graduates of the Library Science and Information Services Department at Central Missouri State University. *Masters Abstracts International, 34*(3), 928. (Publication No. 1377607)

Gooden, A. M. (2001, fall). Citation analysis of chemistry doctoral dissertations: An Ohio State University case study. *Issues in Science and Technology Librarianship*.

Green, R., & Bowser, M. (2003). Evolution of the thesis literature review: A faculty-librarian partnership to guide off-campus graduate research and writing. *ACRL Eleventh National Conference*.

Haycock, L. A. (2004). Citation analysis of educational dissertations for collection development. *Library Resources and Technical Services, 48*(2), 102-106.

Herubel, J.-P. V. M. (1999). Philosophy dissertation bibliographies and citations in serials evaluation. *The Serials Librarians, 20*(2/3), 65-73.

Hovde, K. (2000). Check the citation: Library instruction and student paper bibliographies. *Research Strategies, 17*, 3-9.

Hurst, S., & Leonard, J. (2005, April). *Putting the "B" into BI: An exploratory study of the effect of library instruction on the number, variety, and sources of citations found in business students' term papers.* A poster session at the 12th National ACRL: Currents and Convergence: Navigating the Rivers of Change, Minneapolis, MN.

Iya, J. A. (1996). A citation study of education dissertations at the University of Maiduguri, Nigeria. *African Journal of Library, Archives, and Information Science, 6*(2), 129-132.

Kohl, D. F., & Wilson, L. A. (1986). Effectiveness of course-integrated bibliographic instruction in improving coursework. *RQ, 26*, 206-211.

Kuyper- Rushing, L. (1999). Identifying uniform core journal titles for music libraries: A dissertation citation study. *College & Research Libraries, 60*(2), 153-163.

Magrill, R. M., & St. Clair, G. (1990). Undergraduate term paper citation patterns by disciplines and level of course. *Collection Management, 12,* 25-56.

Malone, D., & Videon, C. (1997). Assessing undergraduate use of electronic resources: A quantitative analysis of works cited. *Research Strategies, 15*(3), 151-158.

Mercer, L. S. (2000, winter). Measuring the use and value of electronic journals and books. *Issues in Science and Technology Librarianship.* Retrieved July 10, 2004, from Expanded Academic Index database.

Okiy, R.B. (2003). A citation analysis of education dissertations at the Delta State University, Abraka, Nigeria. Collection Building, 22(4), 158-161.

Oppenheim, C., & Smith, R. (2001). Student citation practices in an Information Science Department. *Education for Information, 19*, 299-323.

Potter, W. G. (1981). Introduction. *Library Trends, 10*(1), 151.

Radhakrishna, R. B. (1994, December). *A study of core journals used by agricultural and extension educators.* Paper presented at the National Agricultural Education Research Meeting, Dallas, Texas. (ERIC Document Reproduction Service No. ED380550) Retrieved March 22, 2004, from ERIC E-Subscribe database.

Ratteray, O. M. (2002). Information literacy in self-study and accreditation. *The Journal of Academic Librarianship, 28*, 368-375. Rogers, S. A. (2001). Electronic journal usage at Ohio State University. *College & Research Libraries, 62*(1), 25-34.

Sylvia, M. J. (1998). Citation analysis as an unobtrusive method for journal collection evaluation using psychology student research bibliographies. *Collection Building, 17*(1), 20-28. Retrieved October 12, 2003, from ProQuest database.

Thomas, J. (2000). Never enough: Graduate student use of journals – citation analysis of social work theses. *Behavioral & Social Sciences Librarian, 19*(1), 1-16.

Tunon, J. (1999). *Integrating bibliographic instruction for distance education doctoral students into the Child and Youth Studies Program at Nova Southeastern University.* Unpublished doctoral practicum, Nova Southeastern University, Fort Lauderdale, FL. (ERIC Document Reproduction Service No. ED440639)

Young, V. E, & Ackerson, L. G. (1995). Evaluation of student research paper bibliographies: Refining evaluation criteria. *Research Strategies, 13*(2), 80-93.

Cluster ID number_____ **Rater**_____

Subjective Rubric for Doctoral Reference List Resources

Criteria	Level 1 Inadequate	Level 2 Marginally adequate	Level 3 Adequate	Level 4 Superior
Breadth of resources* - number of citations - variety of resources cited	Student used a limited number and/or variety of resources available on topic/Did not show awareness of specialized sources.	Limited number and variety of sources cited	Reasonable number and variety of sources used for topic	Exhaustive search that utilizes a comprehensive number and a full range of types of sources available for topic
Depth: understanding as demonstrated through the citing of historical, theoretical background resources.	Depth of understanding undeveloped by a lack of citations from historical, theoretical background resources.	Depth of understanding emerging as demonstrated through the citation of a limited number of historical, theoretical background resources.	Depth of understanding developed as demonstrated through the citation of a substantial number of historical, theoretical background resources available for the topic.	Depth of understanding exemplary as demonstrated through the exhaustive citation of historical, theoretical background resources available for the topic.
Depth: scholarliness (quality of resources*) - primary resources - empirical research - peer-reviewed - seminal/landmark studies.	Majority of resources superficial/weak	Limited number of scholarly, peer reviewed, resources/too few empirical reviews-superficial	Majority of resources were scholarly, peer reviewed and reasonable no. of empirical research studies	A rich representation of quality, peer reviewed empirical research resources/ very scholarly
Currency* – Criteria take into consideration the availability of resources on the specific topic being researched	Not current – Majority of references older than 10 years from date of dissertation completion	A disproportionate number of unnecessarily dated resources (majority over 5 years)	The majority of the resources published 5 years or less from completion of dissertation	Extremely current – majority of references within 3 years of dissertation completion
Relevancy to the topic	Majority of sources do not relate/pertain to topic	A disproportionate number of sources do not relate/pertain to the topic	Sources generally support/pertain to the topic	Sources directly on target and support/pertain to topic

* Criteria take into consideration the availability of resources on the specific topic being researched

OVERALL SCORE_____**/20**

Appendix B

Sample of Subjective Cover Sheet

CYS ID No: **10802**

Total number of citations in dissertation: **75**

Scoring citations

Number of periodical article citations: **45**

Number of peer-reviewed articles: **39**

Number of academic/scholarly journal citations: **24**

Number of books or book chapter citations: **25**

Number of dissertations: **0**

Number of practicums or theses, etc.: **0**

Number of conference papers: **0**

Number of laws: **0**

Number of government docs/reports/etc.: **2**

Number of ERIC ED documents: **0**

Non-scoring citations

Number of Web sites: **1**

Number of Newspaper articles: **0**

Number of Miscellaneous citations: **2**

Time period break-downs

Total number of citations (all types of docs) published in last three yrs.: **42**

Total number of citations (all types of docs) more than 3 yrs/less than 11 yrs.: **26**

Total number of citations (all types of docs) more than 11 yrs or not dated: **7**

Appendix C

Sample of an Objective Rubric Score Sheet

CYS ID No: 9907

	Total	Total Score	3 Year Total	3 Year Score	10 Year Total	10 Year Score	10 + Total	10 + Score
Citations	141	152.4	21	25	99	108.6	21	18.8
Periodicals	107	125.3	17	23.5	80	91	10	10.8
Books	24	15.7	3	1.5	11	8.2	10	6
Gov. Docs.	1	1.2	0	0	1	1.2	0	0
Conf. Papers	5	6	0	0	5	6	0	0
Diss.	2	4.2	0	0	1	2.2	1	2
Thesis	0	0	0	0	0	0	0	0
Laws	0	0	0	0	0	0	0	0
ED Docs.	0	0	0	0	0	0	0	0
Non-Scoring								
Web Site	0	0	0	0	0	0	0	0
Newspapers	0	0	0	0	0	0	0	0
Misc.	2	0	1	0	1	0	0	0