Improving the Quality of University Libraries through Citation Mining and Analysis Using Two New Dissertation Bibliometric Assessment Tools

Paper presented at the 71st IFLA
(International Federation of Libraries and Associations) Conference, Oslo, Norway
August 22, 2005

Johanna Tuñón tunon@nsu.nova.edu
and Bruce Brydges brydgesb@nsu.nova.edu
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 both 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 143 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 mechanically awarded points for currency, type of document, and certain document-specific criteria, while the second rubric employed a subjective assessment based on the judgment of two assessors using 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 a four-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.

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,* 2003, p. 76*)* 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 contaminated 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 (Gooden, 2001; Herubel, 1991; Kuyper-Rushing, 1999). Only a few studies using citation analysis of dissertation or theses reference lists have focused on the social sciences in general (Sylvia, 1998; Thomas, 2000) or education in particular (Beile, Boote, & Killingsworth, 2003, 2004; Budd, 1988; Haycock, 2004; Iya, 1996; Okiy, 2003).

Some studies have taken citation analysis to the next level to help examine the quality of reference lists, usually in conjunction with subjective rubrics. Rubrics are frequently used as assessment tools to document library research skills. A rubric is defined as a qualitative tool that uses agreed-upon standards in a user-friendly tool to assess individuals' ability to address expected outcomes. Rogers and Graham (1998) defined a rubric as any established set of statements (criteria) that clearly, precisely, accurately, and thoroughly describe the varying, distinguishable, quality (or developmental) levels that may exist in something (a product, organization, creation, system, etcetera). Currently, the most common uses for rubrics in education are evaluating, scoring, and assessing "student work" in order to accurately determine the work's level of quality. The authors' goal was to provide criteria that were based on the best

thinking in the field about what constitutes good performance in order to articulate the essence of what assessors look for when they judge quality (Arter & McTighe, 2001).

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 (Ackerson, Howard, & Young, 1991; Dykeman & King, 1983; Hurst & Leonard, 2005; Kohl & Wilson, 1986; Young & Ackerson, 1995) 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 theses in the field of education 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 the quality of reference lists because they are supposed to be 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). Although a recent study by Beile, Boote, and Killingsworth in two articles (2003, 2004) raised questions about the thoroughness of literature reviews provided in dissertations, these capstone projects are still important to use in this kind of study 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. 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 Doctorate of Education (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 the 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 the quality of 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?

- Do 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 both 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 dissertations of a group of recently graduated doctoral students in Child and Youth Studies (CYS). 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: (See appendix A)

- breadth of resources that considered the number of citations and the variety of citations cited,
- 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 little success in identifying relevant studies. A review of the literature identified a number of studies that used citation analyses to count the frequency with which 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 citation 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; 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, (9) Web sites, and (10) 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, accrediting standards, 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 of how points were mechanically assigned. (e.g. If a student cited and referenced a dissertation, they would receive 2 points. If the dissertation cited was 4 years old at the time the student completed their own dissertation, they would receive an additional 0.2 points because the currency is less than 10 years. Total points for using this dissertation would be 2.2 points. The maximum that any student could get for citing a dissertation of a currency of 3 years or less would be 2.5 points.)

Table 1 Objective Rubric Scoring Scale

	points	**Less than 3yrs old**	**Less than 10 yrs**	**Max pts**
Periodicals (magazines, trade journals)*	0	.3	.2	**.5**
Scholarly periodicals*	1.5	.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	.2	**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**
Dissertations * (published and unpublished)	2	.3	.2	**2.5**
Theses/practicums/action-based research*	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 reference lists rather than a database required a great deal of work. Reference citations were scanned, and the citations normalized and sorted in an Excel software spreadsheet. The citations were then categorized by type. (See Table 1.) When resources fitted into more than one category, the researchers decided to 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 software. A cover page with basic data about the numbers 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 scoring the subjective rubrics. The data was also used to produce the objective scores but only after all reference lists had been scored using the subjective rubric. A total of 10,029 citations in 143 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 all of the reference lists were scored using the two rubrics, the researchers analyzed a number of things. Inter-rater reliability was assessed, and the correlation was found to be quite high, due in part to the fact that the raters worked hard to establish consensus on the descriptors used in each level of the rubric. 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 also existed between the library specialist’s and faculty subject specialist’s assessments using the subjective rubric and the mechanical objective rubric’s totals. (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

The dissertation reference lists 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.

Profiling the typical dissertation reference list

The 10,023 citations were analyzed to get a profile of the types of resources used by doctoral students. A total of 10,023 citations from the 143 dissertation reference lists were analyzed. Since all dissertations available had been included in the sample, there was no sampling error. As Table 3 demonstrates, the resources cited 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 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 the sample for this study were significantly different from the findings in the studies by Beile, Boote, and Killingsworth (2003, 2004) 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 students in this study raised questions about the reasons for these differences. The investigators felt that further research was needed to establish whether this was specific to Child and Youth Study students at NSU or was typical of all Ed.D. students in the university.

Table 3 Citations by Document Type

	Total Citations	Percent of total
Periodical articles*	6774	67.58%
Books	1855	18.51%
Reports*	496	4.95%
Conference proceedings*	113	1.13%
Dissertations/theses/practicums*	76	.76%
Legal documents and laws	30	.29%
ERIC ED documents*	164	1.63%
Newspaper articles*	103	1.02%
Web sites*	177	1.76%
Miscellaneous	235	2.34%
Total	10023	99.97%

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

Table 3 provided an overview of the types of resources cited in all 143 lists. Of the typical 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 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 before

the date of the dissertation's completion, 54.1% were from documents published between 3 and 10 years before completion date of their dissertation, and 20.9% were from documents published 11 or more years before the completion date of the dissertation. Only 53 students out of 143 in the study cited one or more conference papers, and 23 students cited one or more dissertations. The total number of citations per dissertation 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 investigators also disaggregated the 143 reference lists into two cohorts for analysis based on the types of library instruction received. Students in Cohort #1 (n = 69) received a three-step training process (Tunon, 1999) while students in Cohort #2 (n = 74) received the traditional "one-shot" library training session early in the program that lasted 1.5 hours. The researchers compiled the reference lists of the two intervention cohorts (three-step and one-shot training) as well as the dates when the dissertations were completed and approved. (Table 4)

Analysis of data for the different methods of library training indicated that there was no statistically significant difference between the two methods of library training. (Table 5). 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 journals by the late 1990s.

Table 4 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

A Comparison of Distance and Local Cohorts:

The researchers also analyzed differences between the reference lists of students who attended classes locally, with 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 143 dissertation reference lists by dividing the students into local (N=51) and field-based (N=93) cohorts. The researchers found no significant difference between students in local and field-based sites.

Table 5 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%

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 6.)

Table 6 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 digitized and sorted already, the bibliometric data from journal citations was "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, 2004; 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 in the area of child and youth and/or education. All but one of the 20 most frequently cited journals listed in Table 7 were classified as education resources while only 4 were classified in Ulrichs as being on the topic of children and youth, and 1, *School Psychology Review*, was not included in either category in spite of the title. Only 15 of the top 20 titles were categorized by *Ulrich's* as journals, and 16 were identified as being peer reviewed. It should be noted that two of peer-reviewed titles were also classified as magazines. *Ulrich's Directory of Periodicals'* classification scheme provided one standard source for rating periodicals, but the two titles also highlight the problem with Ulrich's use of categories as self-reported by the publishers of those periodicals.

Table 7 Ulrich's Characteristics of Most Frequently Cited Periodicals

	Journal/ Magazine	Academic/ Scholarly	Peer-Reviewed	Education Child/Youth		Number of citations
Educational Leadership	j	a/s		-	ed	317
Phi Delta Kappan	j	a/s		-	ed	163
Exceptional Children	mag	a/s		pr	ed/cy	142
Reading Teacher	j/mag	a/s		pr	ed	104
Intervention in School & Clinic	j	a/s		pr	ed	100
Clearing House	j	a/s		pr	ed	99
Remedial and Special Education	j/mag	a/s		pr	ed	94
Journal of Learning Disabilities	j	a/s		pr	ed	93
NASSP Bulletin	j	a/s		-	ed	84
Adolescence	j	a/s		pr	ed/cy	79
Journal of Educational Research	j	a/s		pr	ed	77
Education	mag	-		pr	ed	68
Young Children	j	a/s		pr	ed/cy	67
Preventing School Failure	j	a/s		pr	ed	66
Educational Digest	mag	a/s		-	ed	63
Journal of Educational Psy.	j	a/s		pr	ed	60
School Psychology Review	j	a/s		pr	psy	56
Journal of Special Education	j	a/s		pr	ed	56
Childhood Education	j	a/s		pr	ed/cy	52
Theory into Practice	j	a/s		pr	ed	52

As with data mining, the structure and format of content really mattered when the researchers used citation analysis for more than simple tools for assessing a journal collection.

Because the migration of resources from print to online in the 1990s, the investigators had hoped to use 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 collected on whether the resources had been retrieved online in order to examine the impact of the Web on the types of resources being cited. The researchers grouped the citations identified as retrieved online by type to see if students were using the Web to retrieve scholarly resources or not.

Table 8 Where Resources Were Retrieved

Type of Resource	Database	URLs	ERIC ED #s	Totals
Journal articles	360	52	0	412
Books	1	1	0	2
Reports	45	49	106	200
Conference papers	0	5	18	23
Dissertations/practicums/theses	5	1	10	17
ERIC documents	5	17	2	24
Laws/legal cases	0	4	0	4
Newspaper articles	6	25	0	31
Web sites	0	179	0	179
Totals	**422**	**333**	**136**	**892**

Table 8 documents students were retrieving a variety of types of scholarly resources retrieved online including journal articles, dissertations, conference papers, reports, and ERIC documents. Journal articles were cited most frequently. Of particular note was the fact that less than half of one percent of the 6774 journal article citations included retrieval statements. A total of 412 journal articles were cited as being retrieved from 158 unique journal titles in either databases or one the Web with an average of 2.60 articles per title. Also notable was the fact that open access journals seemed to have little impact on the research being pursued by these students. In spite of the free online access to open access journals, the CYS students only cited four journals (*Current Issues in Education, Psicologia, Canadian Journal of Psychiatry,* and *New Ideas in Psychology: An International Journal of Innovative Theory in Psychology*), and only one of the four titles were cited as retrieved online. The majority indicated that they retrieved the citations from publisher Web sites or subscription databases. When the proper name of the database was provided, a majority were retrieved from the full-text databases subscribed to by the library (136 from Wilson, 51 from ProQuest, and 36 from InfoTrac).

The lack of retrieval statements made the reliability of data on the use of electronic resources particularly suspect. Although all 143 of the CYS students had been introduced to the Wilson Web Education Full Text database during library instruction sessions, only 36 out 143 students used any retrieval statements. These findings were in line with several previous studies (Beile, Boote, & Killingsworth, 2003; Gooden, 2001; Malone & Videon, 1997) that also questioned students' lack of use of retrieval statements. If students were indeed not using citing resources retrieved online correctly for whatever reason, this type of citing error would have the effect of biasing any conclusions that could be made about the impact of electronic resources on student usage patterns.

ERIC documents made available through the ERIC Document Reproduction Service (EDRS) fall into the category of gray literature not available through journals or book publishers. As Table 8 demonstrates, a number of document types had ERIC ED numbers and most were therefore available through the ERIC Document Reproduction Service (EDRS). Even though the majority of these were in the public domain and therefore had been made available full text online, the citations rarely included retrieval statements. Student problems with formatting resources retrieved from ERIC also presented problems. Of the 496 items categorized as reports, 204 were available as ERIC EDs. However, only 48 of the citations included the ED numbers. In addition, 139 of the ERIC documents where Level 1 documents and therefore in the public domain. EDRS had made Level 1 documents published since 1993 available in a PDF format, but only about a quarter (32 records) of the Level 1 documents included URLS to the ERIC PDF files.

More ephemeral sources like Web sites and miscellaneous resources accounted for only 4% of all resources used. When these types of resources were accessed via the Web, the majority of resources were of a scholarly nature. Government census statistics, unpublished university documents, and reports from various types of agencies were typically cited. Miscellaneous items proved to be a "catch all" for everything from tests, unpublished manuscripts, videos, CDs, and computer software to raw data, poster sessions, PowerPoint presentations, brochures, and more.

Whatever the source of the problem, the widespread lack of inclusion of retrieval statements was disappointing for the 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. As a result, the fact that 99% of the citations (663 out of 10,023 citations) were not formatted as having been retrieved full-text must be treated as suspect. The lack of consistent use of retrieval statements in citations means that they cannot be directly used for insights into patterns of retrieval. In addition, errors caused inaccurate bibliometric information in the citations point to another limitation of using citation analysis for bibliometric data. However, this was less of a problem because the impact of any random citing errors on citing patterns were likely to be cancelled out because of the large data sample.

Conclusions

Citation analysis used in conjunction with either subjective or objective 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 assessment 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 of 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.

Beile, P. M., Boote, D. N., & Killingsworth, E. K. (2004). A microscope or a mirror?: A question of study validity regarding the use of dissertation citation analysis for evaluating research collections. *The Journal of Academic Librarianship, 30*(5), 347-353.

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.

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.

Citation analysis. (2003). In J. Feather & P. Sturges (Eds.), *International encyclopedia of information and library science* (p. 76-78). New York: Routledge.

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.

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. (1991). Philosophy dissertation bibliographies and citations in serials evaluation. *The Serials Librarians, 20*(2/3), 65-73.

Hovde, K. (1999). 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.

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

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.

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

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.

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.

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.

Zuniga, R. (2004). *ACRL/TLT Online Seminar* [Webcast discussion].

Appendix
Brydges/Tuñón Subjective Rubric

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**