

# Understanding Online Journal Usage: A Statistical Analysis of Citation and Use

**John D. McDonald**

*California Institute of Technology, Millikan Library 1-3, Pasadena, CA 91125*

*Email: jmcDonald@library.caltech.edu*

**This study examined the relationship between print journal use, online journal use, and online journal discovery tools with local journal citations. Local use measures were collected from 1997 to 2004 and negative binomial regression models were designed to test the effect that local use, online availability, and access enhancements have on citation behaviors of academic research authors. Models are proposed and tested to determine whether multiple locally recorded usage measures can predict citations and if locally controlled access enhancements influence citation. The regression results indicated that print journal use was a significant predictor of local journal citations prior to the adoption of online journals. Publisher-provided and locally recorded online journal use measures were also significant predictors of local citations. Online availability of a journal was found to significantly increase local citations and for some disciplines, a new access tool like an OpenURL resolver significantly impacts citations and publisher provided journal usage measures.**

## Introduction

Libraries continue to undergo tremendous change due to the increases in the amount of online information produced and used by academic researchers. University libraries no longer focus their efforts as repositories for printed materials but instead are delivering information in print and online formats for most information that they provide their community. The increased availability and accessibility of online journals has been a major improvement in the dissemination of scholarly information. These resources can be accessed any time of day, from anywhere the researcher is located, and remain at that location for simultaneous or future use.

This increase in accessibility has the potential to allow researchers to review a larger amount and greater diversity of material or more time reviewing materials than in the past.

Citation analysis is the standard methodology for studying journal use in information science. Since the Institute for Scientific Information published the first citation indexes (Garfield, 1976), researchers in information science have developed numerous methods for studying how, why, and how often authors cite research articles. Citation analysis is a methodology that has many proponents who view it as a vital research area within the field. Many believe it is a valuable tool that allows librarians to evaluate journal quality, researcher productivity, and journal use. In fact, it is one of the few methods developed in library and information science to be used widely in other fields (Borgman & Furner, 2002).

Detractors point out the inherent faults associated with citation analysis, including the highly individualized nature of citing material (Bonzi & Snyder 1991; Brooks 1986; Cano 1989; Cole & Cole 1972; Gilbert 1977; MacRoberts & MacRoberts, 1987, 1988, 1989; Vinkler 1987), the differences between fields in citation culture and journal content or format (Seglen, 1997a, 1997b, 1998), the tendency for well known authors to be cited exponentially (Merton, 1973), and problems associated with the collection and indexing of articles by ISI. To avoid these problems, researchers have tried to develop other citation-based methods of ranking journals (Nisonger, 1999). Borgman and Furner (2002) provide a thorough review of citation analysis and bibliometrics. Despite the fact that the nature of citation is still not fully understood, citations remain a highly quantified, easily collected, easily obtained, and easily manipulated set of data and has been considered to be a form of library use (Kelland & Young, 1994).

Usage analysis, in contrast to citation analysis, is an emerging area of bibliometric research. Issues surrounding analysis of online journal usage statistics have lately come to the forefront of library and information studies research (Luther, 2000) and have helped focus the need to standardize the collection and reporting of this metric. As a result, two groups have been formed to develop standards on the collection and

reporting of online journal statistics (NISO, 2002; Project Counter, 2002).

Recent studies have explored the use of online journals but have focused on the use of online-only (or born digital) journals or use of materials by format (i.e. online versus print) (Mercer, 2000; Morse & Clintworth, 2000). Other studies have examined how online materials are cited in specialized online-only journals (Harter, 1998), or preprint archives (Kaplan & Nelson, 2000; Brown, 2001) but rarely on materials that are simultaneously published in print and online formats (Ke et. al., 2002; Tenopir & King, 2002; Rogers, 2001). Another focus of recent research is the practical application of online journal usage statistics to evaluate collection decisions, assist in cost-benefit analysis, and develop long term planning initiatives (Davis, 2002a, 2002b, 2002c; Goodman, 2002; Montgomery, 2000). Bleic (1999) provides one of the only comparisons of the methodologies of online journal use measurement.

Statistical methodologies have been used to explore the nature of online journal usage, the effect on print collections, usage by subject area, and the ability to use citations to predict author impact. Tenopir & King (2004) used descriptive statistics derived from a survey to conclude that faculty in the medical sciences prefer traditional print formats to online. De Groote & Dorsch (2001) used repeated measures ANOVA (analysis of variance) tests to determine that print journal usage and interlibrary loan requests decreased since the introduction of online journals. Feitelson and Yovel (2004) modeled citation accumulation to predict author rankings in computer science. Their study is one of the few attempts to use inferential statistical modeling to predict future events such as the receipt of citations by authors. Shin (2003) used ANOVA to study the effect that conversion to an online formats has on a journal's impact factor. The author found that journals published only in print formats had no significant changes in impact factor, but those that published in print and online formats had a significant increase in impact factor. Tsay (1998), in contrast to Scales (1976), found significant rank-ordered correlations of frequency of use and frequency of citation in life sciences and clinical medicine journals. Antelman (2004) studied Open Access journal articles and found that the free availability of these journals significantly affected citations in four disciplines. This study is especially relevant to the current analysis because she used the Wilcoxon signed-rank test, a statistical test for skewed distributions. Stemper and Jaguszewski (2003) reviewed the differences between local and vendor-supplied counts of online journal usage statistics and found that local usage statistics differed from vendor supplied usage statistics for a large percentage of titles studied. The authors used only descriptive statistics, leaving the statistical significance between the differences still in doubt.

### Purpose

Although citation analysis has a strong theoretical foundation in LIS research, there are few studies linking citations to other journal usage measures. The ability to

monitor and measure the use of online journals provides researchers with new data to explore the relationship between citation and usage. These types of journal usage studies might serve the same purpose that citation studies have served in the past, to examine relationships between research literature, readers, and authors. Since the vast literature on citation analysis and bibliometrics relies on the citation as the most valid measure of the use of research literature, studying the relationship between online journal use and citation will help link previous theories of citation to current studies of online journal usage studies. Usage-based studies incorporating inferential or other advanced statistical analysis, and not solely descriptive statistics, will also help build a solid theoretical foundation for usage-based bibliometrics. Librarians also have an intense interest in the practical utility of online journal usage statistics and especially in developing practical methodologies to analyze these statistics.

Faculty and students typically select materials that are the easiest to identify and access. Joswick and Stierman's (1997) analysis of local citation patterns among faculty, graduate, and undergraduate students showed that all three groups select material available locally or otherwise easily attained over materials that are not locally available. Other research has also shown that faculty and students prefer online materials to print (Brown, 1999; Morse & Clintworth, 2000; Rogers, 2001). The ease of access and increased functionality of online materials allow users to review more material in a shorter amount of time, which may lead them to review additional material, review more diverse material, or produce publications at a faster rate. These changes in information seeking and use behaviors should be evident from new patterns of citation or online journal usage.

A number of studies have recently focused on online journal usage by local user populations. Darmoni and Roussel (2002) proposed a statistic called *Reading Factor* that was defined as the number of full-text uses of an article divided by the total number of articles in the database in which it is found. Juxtaposed against ISI's journal impact factor, which is global in nature, this could be an excellent local quantitative measure when an institution uses a single full-text article database. Other types of local user analyses are beginning to occur more frequently with the provision of online usage statistics through transaction log methods, especially Web server logs and publisher provided usage data (Davis, 2002a; Davis, 2002b; Davis, 2004; Black, 2005; Stemper & Jaguszewski, 2003).

Librarians use journal usage statistics for many practical applications: to begin or end subscriptions, to justify budget allocations, to prioritize research areas, programs, and education, and to seek funding, while University administrations and faculty bodies use citation information for many of the same purposes. Publishers are beginning to price journal and database subscriptions based on the number of articles retrieved from them, giving librarians more reason to develop critical analytical tools regarding online journal usage.

Online journal usage analysis needs to incorporate research methods and predictive models that build upon prior bibliometric theories and research, especially prior

citation analyses. Bridging the previous theoretical literature on citation analysis to newer research on online journal usage is one of the major goals of this study. A better understanding of the relationship between citation and usage will allow researchers to propose and test theories of the relationship of citation to usage. In addition, the current study will explore the relationship between citations and three measures of use, two locally recorded and one provided by journal publishers, explore relationships between citations and online availability, and explore the effect that discovery and access tools have on citation. A secondary purpose of the study is to determine the suitability of inferential statistical methods, specifically negative binomial regression, for citation and online journal usage analysis.

## Data

The dataset includes 1521 journals that are owned California Institute of Technology (Caltech) library, a scientific research university. Information about these journals has been collected from 1997 through 2004 and include variables such as the date of first local online availability, total local use of online (measured by the library), total local use of online (as reported by publisher), total local citations, and total articles published by the university's authors (as reported by ISI). Publisher provided online journal usage reports were not common over the study's time period and limited the dataset in years and total titles included. Specifically, the earliest year of online journal use reported was in 1997 but included only sixteen journals from one publisher. The journals were also collated into nine broad subject categories: astronomy, biology, chemistry, computer science, engineering, general science, geology, mathematics, and physics. It is important to note that although the set includes over 1500 journals, some journals lacked information for certain variables and therefore could not be included in every model or test.

## Definitions

The variable, *Print Use*, is a locally recorded count, cumulated by bar-code scanning materials in reshelfed in the libraries, circulated to patrons, or used for the library photocopy service program. Each time a bar-code is scanned equals one use of a journal. Many combinations of these measures have been utilized in previous print journal usage studies (Langlois & Von Schul, 1973; Millson-Martula, 1988; Chrzastowski & Olesko, 1997; Bleicic, 1999).

The variable, *Access Use*, is a measure of local online journal use derived from transaction log counts of the two main routes that users accessed online journals. One use measure, total annual links followed from the library's alphabetic online journal list was combined with a second measure, total annual OpenURL resolver links by journal. The two measures are mutually exclusive since a user could access online journals through the library web page or the OpenURL resolver, but not through both at the same time. While neither measure indicates that any particular article from any given journal was actually downloaded, printed, or used, they are

roughly the online equivalent of *Print Use*, simply indicating that a user was interested in accessing the particular journal for some reason. Transaction log analysis, especially analysis of search and retrieval counts and search strategies, has a long history as an analytic technique in bibliometrics (Kaske, 1993).

*Publisher Use* is the total annual full-text article downloads, by journal, as reported by the publisher to the library. Publishers' usage reporting mechanisms initially developed idiosyncratically and, depending on the publisher, included a variety of counts, such as: use of abstracts, use by IP address, use by file type (PDF & HTML), and page downloads. Current endeavors by Project Counter and NISO to standardize usage counting and reporting have improved the dissemination and validity of publisher reported statistics. At this point in time, full-text downloads are the most commonly reported measure of article use in the industry despite issues surrounding their consistency and commonality (Luther, 2000).

Local citations, total citations, and total articles published by university authors, were compiled using the ISI Web of Science databases for the appropriate years. Cited references for every article published by university authors were compiled by journal and included every citation to a particular journal in every article published in the given year by a university author.

There are two potential arguments against the appropriateness of these use measures for the evaluation of citation behavior. First, some usage, locally recorded or publisher provided, is not equivalent to all other usage. For the purposes of this study all 'uses' of a journal are assumed to be equal. Accessing a particular journal for any reason, including browsing, vanity searching, inadvertent use, verification use, download of material that was not used, etc. is equivalent to use of a journal for reading, research, writing, and citing. 'Mis-use', 'non-use', or other non-traditional use of material is still a form of 'use' in the practical application of collection management in libraries. More so, there is no indication that a disproportionate amount of non-traditional use occurs by certain user groups, whether subject-based or demographically-based. Second, an argument can be made that the data collection techniques might leave out a significant amount of use from researchers accessing online journals through a variety of other non-traditional methods. Potential journal use that is not captured in the current study might include Web browser bookmarks that bypass local data collection methods, use of other library's subscriptions, personal journal subscriptions, interlibrary loan, local use by non-affiliated researchers, and other article acquisition techniques by individual researchers. Again, there is no previous research to indicate bookmarking, use of personal subscriptions, or other article acquisition techniques that are not measured through the current methods, have significantly different variances by publication format or user group. This research is solely an attempt to examine if the amount of use from the current measurement techniques of local transaction logs (*Access Use*) and publisher reported full-text downloads (*Publisher Use*) predict citations, a use measure that already has a solid theoretical foundation. A finding of non-significance can lead to future research

exploring why, and future hypotheses could focus on non-measurable uses or types of use that do not lead to citation.

## Hypotheses

Five hypotheses are proposed and tested in the current study. First, citations are positively correlated to print journal usage. Previous studies have indicated that usage is at least a partial requirement for citation in a paper and this study attempts to link the pre-online journal usage recorded at Caltech to citations in locally authored articles. A significant relationship will provide a foundation for subsequent testing of the relationship between online journal use and citation. Second, citations are positively correlated to locally measured online journal usage. A significant relationship between print use and citation leads to a logical hypothesis that online use should also have a significant relationship with citation. Third, citations are positively correlated to publisher provided online journal use. Similar to the two previous hypotheses, publisher provided online journal use should also have a significant relationship with citation if use is a requirement for citation. Fourth, online availability increases local journal citations. If online journals are more accessible and useful than traditional print journals, there should be a significant increase in citations after provision of the journal in this format. Fifth, tools that increase discoverability and access to online journals increase local journal citations and increase local usage. While this study only attempts to analyze the effect that increased discoverability has on citation and usage, removing barriers to use, including subscription access restrictions, local access restrictions, and other barriers could lead to measurable increases in usage and citations.

## Methodology

### Section 1: Relationship between citation and use

*Local Citations*, as well as *Access Use*, *Publisher Use*, and *Print Use*, are count data variables ranging from zero events to infinity, with no negative numbers. Furthermore, the variables typically include many zero events and few occurrences of very large events per journal. The dataset is therefore non-normal in its distribution and the independent variables in this study are not independent observations. Therefore, these violations of the fundamental assumptions of the model render ordinary least squares regression inappropriate. An alternate regression model that is commonly used to analyze non-normal count data is Poisson regression (Gardner, Mulvey, Shaw, 1995; Cameron & Trivedi, 1998). Further examination of the dependent variable (*Local Citations* 2004) shows extreme overdispersion with a variance (44297.36) dramatically higher than the mean (42.52). Using a Poisson model with this data may produce incorrect estimates of its variance and incorrectly indicate significance for some of the independent variables. An alternate regression model

that can account for overdispersion is negative binomial regression because it includes a random term that reflects the unexplained variance (Gardner, Mulvey, Shaw, 1995; Cameron & Trivedi, 1998). Although these types of regression models are used frequently in other social sciences fields such as political science, epidemiology and demography, the only example of negative binomial regression in library and information science research is a study by Van Dalen and Henkens (2001) that used negative binomial regression to evaluate citedness of articles.

The first hypothesis tested was whether print usage statistics can predict local citation frequencies. *Print Use* was regressed on *Local Citations* in 2001, limited to journals that were available locally only in print formats before 2001. The dataset exhibits a non-normal distribution, so Kendall's tau-b, a non-parametric correlation, was used to evaluate the suitability of the independent variables since it is more robust to non-normal data and ties in ranked correlations. Table 1 shows the Kendall's tau-b ranked correlations of *Local Citations* received in 2001 and *Print Use* measures from 1998 to 2001. The strong correlations between the *Print Use* variables indicate that it is a reliable measure from year to year and its strong correlation with *Local Citations* indicates that it is good variable to relate to citations.

**TABLE 1: Correlation of Citations 2001 and *Print Use* Measures# (N=458)**

	Local Citations 2001	Print Use 2001	Print Use 2000	Print Use 1999	Print Use 1998
Local Citations 2001	1.00				
Print Use 2001	0.40	1.00			
Print Use 2000	0.43	0.67	1.00		
Print Use 1999	0.44	0.65	0.71	1.00	
Print Use 1998	0.39	0.62	0.66	0.70	1.00

#Kendall's tau-b

The regression model included the independent variables of *Print Use* for 1998 to 2001 and the dependent variable of *Local Citations* in 2001. The model was limited to variables from before 2002, because more journals became available online and impacted print journal usage after that date. Model 1 (Table 2) shows the results of the first negative binomial regression on these variables. Incident rate ratios (IRR) are included rather than untransformed coefficients to help in the interpretation of the model. IRR coefficients

indicate the effect of a one unit change in an independent variable on the dependent variable.

Not only are three of the variables not significant, but the IRR values of less than one for *Print Use* in 2001 and 1998 are cause for concern, since theory dictates that higher *Print Use* results in higher citations, not lower. In addition, the high Kendall tau-b correlation between the *Print Use* variables indicates that a model including all four suffers from multicollinearity, making the highly correlated variables unnecessary. To solve this, a second negative binomial regression model was developed that included just *Print Use* in 1999. The more negative BIC prime statistic indicates that Model 2 is preferred over Model 1.

**TABLE 2: Negative Binomial Regression Models predicting Local Citations from Print Use variables**

Local Citations 2001	Model 1	Model 2
N (journals)	302	302
<i>Incident Rate Ratios (standard errors in parenthesis)</i>		
<i>Print Use</i> in 2001	0.999 (0.003)	
<i>Print Use</i> in 2000	1.000 (0.005)	
<i>Print Use</i> in 1999	1.016 (0.007)**	1.013 (0.002)**
<i>Print Use</i> in 1998	0.998 (0.004)	
$\alpha$	2.979 (0.289)**	2.984 (0.298)**
Contrasts		vs. Model 1
BIC'	-159.72	-176.56
LR chi2	182.56	182.27
Prob > LR	0.000	0.000
McFadden's adj R <sup>2</sup>	0.09	0.10

\*significant at the .05 level  
\*\*significant at the .01 level

These results indicate that *Print Use* predicts *Local Citations* after a delay of two years and that negative binomial regression is a valid regression analysis for this type of data. In addition, the R<sup>2</sup> of 0.10 indicates that the *Print Use* variable is a relatively large portion of a model that predicts the extremely complex act of citation. These results are problematic for libraries since very few libraries have recorded or continue to record usage of their print collection and most usage studies have shown print usage measures have started falling with the proliferation of online journals (De Groote & Dorsch, 2003; Black, 2005). The same situation has occurred at Caltech as well. Table 3 shows the average change in

local usage of the print collection, by subject area, of each journal after the online version became available for each year. Table 4 shows the change in *Print Use* after provision of the online version from two years before to two years after. The Wilcoxon signed-rank test for all subjects (except General Science) was significant and included large z-scores, indicating that the difference in *Print Use* from before a journal was available online to after was not due to random chance. Even though some *Print Use* of the journal collection persists, it is undoubtedly due to use of the non-online portion of the back issues. Exactly what portion of use this accounts for remains unknown since no measures are recorded about the use of each journal by year of publication. Quite simply, this data indicates that provision of a journal in an online format results in fewer uses of print materials.

**TABLE 3: Average change in Print Use by Year and Subject**

Subject	##	2004	2003	2002	2001	2000	1999	1998
Astronomy	20	-10%	8%	-37%	-27%	-24%	-25%	-10%
Biology	280	-34%	-18%	-48%	4%	-10%	-25%	10%
Chemistry	105	-31%	-33%	-52%	47%	-8%	-31%	4%
Computer Science	71	-22%	-20%	-26%	-24%	12%	-28%	-4%
Engineering	188	-20%	-19%	-17%	-21%	-15%	-17%	3%
General Science	18	-55%	-4%	-46%	-10%	-12%	-23%	15%
Geology	144	-33%	-20%	-16%	-3%	2%	-14%	7%
Mathematics	141	-45%	-25%	-17%	13%	7%	-14%	8%
Physics	100	-36%	-2%	-31%	-5%	-12%	-17%	-3%
<b>Total</b>	<b>1067</b>	<b>-34%</b>	<b>-19%</b>	<b>-39%</b>	<b>7%</b>	<b>-8%</b>	<b>-23%</b>	<b>5%</b>

**TABLE 4: Average Print Use Before & After Online Availability**

Subject	Journals	Before	After	Change	Wilcoxon Signed-rank Test#	
					z	P>z
Astronomy	11	172	74	-57%	2.09	0.04*
Biology	183	98	43	-56%	10.64	0.00**
Chemistry	70	144	84	-42%	6.56	0.00**
Computer Science	45	34	14	-59%	5.77	0.00**
Engineering	143	53	24	-55%	8.99	0.00**
General Science	7	691	585	-15%	1.28	0.20
Geology	98	112	52	-54%	6.04	0.00**
Mathematics	87	43	29	-33%	5.65	0.00**
Physics	70	149	79	-47%	5.87	0.00**
<b>Total</b>	<b>714</b>	<b>1496</b>	<b>984</b>	<b>-34%</b>	<b>19.418</b>	<b>0.00**</b>

# Test on *Print Use* 2 yrs before vs. 2 yrs after online availability:  
Ho: *Print Use* Before = *Print Use* After  
\*significant at the .05 level \*\*significant at the .01 level

## Section 2: Relationship between citation and online journal use measures

The confirmation that print journal use is a large and significant variable in predicting local journal citations leads to an examination of whether locally recorded journal use predicts local citations. Local use measures include indications of use from two mutually exclusive transaction (Web server) logs: an online list of titles and a reference linking server. Both measures have been collected since 2001 and were combined into the variable titled *Access Use*. This measure indicates the traceable access route or gateway that a user may take to get to a particular journal included in the study. A library user moves from these two discovery and access resources to a publisher's site in order to acquire a particular article from a journal. The recorded measure is an indication of interest in a particular journal and is roughly equivalent to the print journal usage measures detailed previously. The second measure, also collected for most journals, since at least 2001, is publisher provided usage statistics of full-text article downloads for each journal. While this measure may seem to be the most reliable, publishers have not perfected their technical systems for measuring this use, their reporting systems, and still have a possible economic or business motivations for mis-reporting or over-reporting these statistics.

Table 5 lists the Kendall's tau-b correlations between *Local Citations* in 2004 and both sets of usage measures. The variables are once again highly correlated, indicating that they are reliable and appropriate independent variables. The correlations between both use measures and citations are much greater than those between *Local Citations* and *Print Use*, indicating that they may have captured more of the use leading to citation than print journal usage did previously. Once again, the independent variables are also highly correlated with each other, possibly indicating multicollinearity and significant redundancy.

A negative binomial regression was run on all years of *Publisher Use* and *Access Use* and then repeated for additional models incorporating a variety of independent variables to find the most parsimonious model. These subsequent models included independent variables that paralleled the former *Print Use* regression model and

other combinations of *Publisher Use* and *Access Use* variables to determine if there is a more rapid accumulation of citations due to online journal availability. Table 6 shows the model characteristics for the three negative binomial models.

The results of Model 1 indicated that the inclusion of all years of data for both usage measures would be inappropriate for this model. Three variables were negative -- a surprising result -- while three other variables were only significant due to the inclusion of the other redundant variables (multicollinearity). A revised model that followed the conventions of the *Print Use* model and included both local online use measures from 2002 showed potential since both variables were positive and significant, but a model that included more current measures from 2004 was better, showing a higher  $R^2$  and more negative BIC' statistic (Model 2). Grouping measures for the previous two years into single variables did not improve results significantly, but reducing the model to include just *Publisher Usage* from 2004 resulted in a satisfactory model as well, with positive and significant coefficients and a comparable  $R^2$  to other models. Model 2, which included locally recorded usage data as well as publisher provided usage data from 2004 was the best model for predicting citations in 2004, while Model 3, which used *Publisher Use* 2004 as the only independent variable was the most parsimonious model. Table 7 shows predicted citations, based on Model 2, for given values of *Access Use* and *Publisher Use*. It is important to note that the  $R^2$  of .05 for Model 2 (*Access Use* and *Publisher Use*) or Model 3 (*Publisher Use*) indicates that using either the two variables together or one alone explains 5% of the variability in *Local Citations*. While this does not seem high, it is significant and relatively high for a one or two variable regression model. This indicates that at least one usage measure can be used as one variable in a multi-variable model to predict local citations. Furthermore, since the usage measures are not only highly correlated with local citations and but also each other and a regression model including just one of the measures (i.e. Model 3) is satisfactory, the different use measures are interchangeable and may be used to model missing values of the other measure.

**TABLE 5: Correlations of *Local Citations* and Other Use Measures# (N=639)**

	Citations 2004	Publisher Use 2004	Publisher Use 2003	Publisher Use 2002	Access Use 2004	Access Use 2003	Access Use 2002
Citations 2004	1.00						
Publisher Use 2004	0.62	1.00					
Publisher Use 2003	0.63	0.84	1.00				
Publisher Use 2002	0.60	0.79	0.80	1.00			
Access Use 2004	0.53	0.57	0.57	0.55	1.00		
Access Use 2003	0.59	0.64	0.65	0.63	0.66	1.00	
Access Use 2002	0.64	0.71	0.70	0.72	0.61	0.73	1.00

#Kendall's tau-b

**Table 6: Negative Binomial Regression models predicting citations in 2004 (standard errors in parentheses)****Model 1:** *Publisher Use* and *Access Use*, 6 independent variables**Model 2:** *Publisher Use* and *Access Use* 2004, 2 independent variables**Model 3:** *Publisher Use* 2004, 1 independent variable

Variable	Model 1	Model 2	Model 3
N (journals)	639	639	639
<i>Incident Rate Ratios (standard errors in parentheses)</i>			
Publisher Use 2004	1.001 (0.000)*	1.001 (0.000)*	1.001 (0.000)*
Publisher Use 2003	0.999 (0.000)		
Publisher Use 2002	1.000 (0.000)**		
Access Use 2004	0.999 (0.001)*	1.001 (0.000)**	
Access Use 2003	1.003 (0.001)**		
Access Use 2002	0.999 (0.001)		
$\alpha$	3.278 (0.190)*	3.363 (0.194)	3.445 (0.198)
<b>Contrasts</b>			
BIC'	-233.776	-244.042	-235.247
LR chi2	272.535	256.962	241.707
Prob > LR	0.000	0.000	0.000
McFadden's R2	0.057	0.052	0.050

\*significant at the .05 level

\*\*significant at the .01 level

**Section 3: Relationship between citation & online availability**

Citation is clearly related to usage and usage measures are one variable that can contribute to a model to predict citations. The regression models outlined previously indicate that *Print Use* could predict citations prior to the introduction of online journals, and either *Publisher Use* or *Access Use* can predict *Local Citations* of journals. Another phenomenon apparent from the descriptive statistics of this study is that usage and citation are both increasing in recent years. Online journal usage is increasing, as measured both locally and by publishers. Likewise, local citations are increasing as well, but at a much lower rate.

**Table 7: Predicted Local Citation rates for Model 3**

Access Use 2004	Publisher Use 2004	Predicted Local Citations 2004
10	10	8
50	50	9
100	100	10
200	200	11
300	300	13
400	400	14
500	500	18
1000	1000	43
2000	2000	225
3000	3000	1424

Previous research has indicated that one of the primary factors that influence usage of journals is ease of access. Both online availability and tools that assist users in discovering and accessing online journals lead to increased usage and subsequently increased citation counts. Building a truly experimental model for testing these hypotheses through an analysis of citations of online versus non-online materials is difficult since 1) the version format (online or print) of a typical citation can not be determined for most citations previously reported, 2) most journals have online versions that are already heavily preferred by researchers, and 3) journals that still have no online version are the most infrequently used and cited. Table 8 shows the local citation rate per journal for each format from 1997 to 2002. Citation rates for both print only and print plus online journals are increasing in recent years. The large difference between the two formats is due to the low citations of the particular journals that make up the print only group, since most highly cited journals have been offered in online versions since at least 2001 or much earlier. The large jump in citations from 1999 to 2000 and from 2000 to 2001 are probably caused by the conversion of most of the journals in the print plus online category during those time periods. Table 9 shows the *Local Citations* per journal by the date the online version was added. The years, 1997, 1999, 2003, and 2004, have very few journals included and can therefore be skewed by journal subject or type. Those years that did have large numbers of journals that added online versions clearly show that *Local Citations* in the year after the journal became available online were substantially higher than the average citations the year before the online availability.

**TABLE 8: Average Local Citations per Journal by Format**

Format	#	1997	1998	1999	2000	2001	2002	2003	2004
Print Only	119	2.96	3.05	2.91	3.71	4.20	4.19	3.92	3.82
Print plus Online	1317	31.65	31.96	37.01	39.24	45.25	44.76	46.04	42.56

**TABLE 9: Average Citations per Journal by Date Online Added**

Date Online Added	#	Average Citations Per Journal								
		1997	1998	1999	2000	2001	2002	2003	2004	
1997	34	254.50	249.21	354.29	329.97	378.15	370.74	400.65	338.53	
1998	183	<b>70.19</b>	67.19	<b>73.78</b>	80.20	89.57	89.43	91.16	82.77	
1999	86	30.80	<b>32.28</b>	38.43	<b>40.42</b>	44.01	42.64	44.02	41.73	
2000	243	23.88	27.44	<b>29.74</b>	33.47	<b>48.52</b>	47.75	45.52	44.30	
2001	530	10.74	11.18	12.29	<b>13.99</b>	14.91	<b>14.94</b>	14.65	13.99	
2002	210	25.92	26.28	26.65	29.28	<b>29.47</b>	29.09	<b>32.25</b>	31.27	
2003	79	5.44	3.67	4.44	4.16	6.14	<b>6.10</b>	6.27	<b>6.28</b>	
2004	20	2.00	1.95	1.20	1.85	4.10	4.05	3.25	2.90	

In summary, print use declines after a journal becomes available online. In addition, online journals are cited at a higher rate than non-online journals, and citations increase after a journal becomes available online despite publishing, ostensibly, the same quality of material. Table 10 presents the means and standard deviations for local citations, by discipline, for the years preceding and following the provision of it in an online format. The Wilcoxon Signed-Rank test indicates that most disciplines have a significant difference in journal citations between two years before and two years after being provided online. Two years was chosen to offset any effects that user education, acceptance of online journals, or the time lag inherent in usage-to-citation appearance might have on the data. The mean citations are increasing for most disciplines, but the standard deviations are also increasing, indicating that more heavily cited journals are gathering more citations while

**TABLE 10: Citation Rates Before & After Online Availability**

Subjects	Journals	2 Yrs Before		1 Yr Before		Yr Online		1 Yr After		2 Yrs After		Wilcoxon Signed-rank Test#	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	z	P>z
Astronomy	13	313	390	702	1320	812	1592	914	1461	1156	2357	-2.3	0.02*
Biology	337	17	60	21	69	21	65	21	61	25	74	-4.3	0.00**
Chemistry	115	33	123	34	117	33	108	32	89	33	94	-1.3	0.18
Computer Science	83	7	14	6	12	7	18	8	23	8	21	0.7	0.48
Engineering	147	10	21	11	26	10	25	11	26	12	26	-1.3	0.19
General Science	10	326	670	392	688	418	752	424	780	452	809	-0.7	0.49
Geology	112	19	42	20	40	25	49	24	50	25	48	-4.2	0.00**
Mathematics	98	4	5	6	8	5	7	6	10	7	11	-3.8	0.00**
Physics	100	64	160	67	148	67	157	86	270	89	266	-2.9	0.00**
<b>Total</b>	<b>1042</b>	<b>27</b>	<b>115</b>	<b>35</b>	<b>196</b>	<b>37</b>	<b>225</b>	<b>40</b>	<b>230</b>	<b>46</b>	<b>318</b>	<b>-7.3</b>	<b>0.00**</b>

# Test on citations 2 years before and 2 years after online provision (Test  $H_0$ : Cites Before = Cites After)

\*significant at the .05 level    \*\*significant at the .01 level

citations to less cited journals continue to decline. The negative and significant z-scores on the Wilcoxon Signed-Rank test for five of the subject categories indicate that more journals increased in citation than decreased during the two periods of comparison. Further study is needed to understand why journals in the other four categories (chemistry, computer science, engineering, general science) were not significant in their citation increases after online provision and potential reasons include the low number of journals in the category (i.e. general science), the nature of publication and citation for the discipline, the coverage of journals by ISI in the subject area, or the maturity of the discipline in using online materials.

#### Section 4: Relationship between citation & access tools

Increased usage and online availability, as indicated in sections 1 and 3 respectively, lead to increased citations. Online journals' ease of use and speed of access allows authors to review more materials or produce publications more quickly that result in increases in citation counts. Access tools that reduce barriers to use and improve the identification and acquisition of relevant journal literature should have a similar significant impact on citations. These access tools could include any technical, social, or economic mechanism that makes it easier for the researcher to discover, access, and use material. Examples of access tools could range from the purchase of an abstracting and indexing database, improved interfaces and retrieval mechanisms for existing databases, bibliographic instruction courses for researchers, or new tools to organize and present information to the researcher.



Caltech introduced an OpenURL resolver as a major new access tool in 2001 that provides a good basis for testing this hypothesis. The resolver is an integrated reference linking server that connects bibliographic databases with full-text journal articles through one-click access. Barriers to effective use of journals have included intellectual barriers, such as knowing how to search databases and where to locate materials in physical locations, as well as organizational barriers, such as inadequate bibliographic cataloging and changing Internet locations for the materials. The resolver removes most of these access barriers by directly connecting users to articles and improving information about the library's collections (i.e. notifying a user when full-text is or is not available).

Table 11 shows the means, standard deviations, and Wilcoxon Signed-Rank test scores by discipline for journals that were already available online by 2001. The Wilcoxon Signed-Rank test was done on *Local Citations* for the year before the introduction of an online version compared to 2003, two years after the introduction of the resolver. The negative z-scores indicate that citations after the introduction of the resolver increased compared to citations received before the journal became available online and the significant tests for five of the nine disciplines indicate that access tools may have different effects by discipline. The negative and significant z-scores for astronomy, biology, chemistry, mathematics, and physics indicate that those disciplines had more journals increase in total *Local Citations* and the change could not be due to random chance, while the other four

disciplines did not have significant scores. It is possible that the significance would change with different sample journals, a different potential set of users, or with more journals in the dataset for some disciplines. This suggests that there are real differences by subject discipline in their citation behavior, their use of the OpenURL resolver, or a combination of the two behaviors.

Even though there may not yet be overwhelming evidence to indicate that access tools like OpenURL resolvers have an effect on citations, resolvers do have a large and significant effect on *Publisher Use*. Table 12 shows the *Publisher Use* for 2000 through 2004 and the Wilcoxon-Signed Rank test scores by discipline. The large and significant z-scores for the whole set of journals, biology journals only, and also all non-biology titles indicates that *Publisher Use* increased after introduction of the resolver. Some subject groupings z-scores were not significant but the small sample sizes for those disciplines make those results questionable. Most raw counts of online usage have increased over time and probably have regardless of the provision of an access tool like an OpenURL resolver. These increases may be due to increased general awareness of online journals by users, increased acceptance of using them, increased utility of the online version, changing information seeking, storing, and retrieval behaviors of users, or other changes that are not directly measurable. But undoubtedly, access tools such as OpenURL resolvers result in increased publisher reported usage.

**TABLE 11: Citation Rates Before & After Introduction OpenURL Resolver (only journals that were available online prior to 2001)**

Subjects	Journals	2 Yrs Before		Yr Online		2 Yrs After		Wilcoxon Signed-rank Test#	
		Mean	SD	Mean	SD	Mean	SD	z	P>z
Astronomy	6	2187	3380	2443	3601	2659	3906	-2.1	0.04*
Biology	109	49	103	53	102	52	102	-2.7	0.00**
Chemistry	37	57	145	83	195	80	184	-2.8	0.00**
Computer Science	25	3	5	3	5	4	5	-1.8	0.07
Engineering	25	11	19	15	26	14	26	0.01	0.99
General Science	6	778	926	832	992	873	1032	-1.1	0.29
Geology	21	26	54	28	49	29	49	-1.4	0.16
Mathematics	33	6	9	7	9	10	11	-2.6	0.01**
Physics	41	62	127	80	166	71	146	-2.4	0.02*
<b>Total</b>	<b>303</b>	<b>95</b>	<b>558</b>	<b>109</b>	<b>605</b>	<b>113</b>	<b>653</b>	<b>-5.8</b>	<b>0.00**</b>

# Ho: Cites After Online availability = Cites After OpenURL resolver

\*significant at .05 level

\*\*significant at .01 level

**TABLE 12: Publisher Use Before & After Introduction of OpenURL Resolver (only journals that were available online by 2000)**

Subjects	Journals	Publisher Use Year Before Resolver		Publisher Use 2001		Publisher Use Year After Resolver		Wilcoxon Signed-rank Test#	
		Mean	SD	Mean	SD	Mean	SD	z	P>z
Astronomy	1	347	0	813	0	1408	0	-1.0	0.32
Biology	104	638	1625	847	2079	957	2351	-5.9	0.00**
Chemistry	42	1388	3248	1553	3889	2542	7294	-4.9	0.00**
Computer Science	14	197	429	224	490	175	239	-1.6	0.10
Engineering	20	92	200	164	310	174	312	-2.4	0.02*
General Science	3	16243	15571	20938	20345	26553	26506	-1.4	0.17
Geology	22	46	183	44	143	144	374	-3.1	0.00**
Mathematics	29	61	158	82	155	123	185	-3.6	0.00**
Physics	28	217	324	1189	2193	1684	2048	-4.0	0.00**
<b>Total</b>	<b>263</b>	<b>711</b>	<b>2749</b>	<b>988</b>	<b>3552</b>	<b>1320</b>	<b>4988</b>	<b>-10.4</b>	<b>0.00**</b>

# Ho: Publisher Use 2000 (before resolver) = Publisher Use 2002 (after resolver)

\*significant at .05 level

\*\*significant at .01 level

## Conclusions

This study indicates that although citations remain a difficult phenomenon to understand, some of the variability can be explained by local usage measures. Print journal usage could have been an important variable to include in a predictive model for local citations, but the provision of online journals has rendered this variable unusable. Locally collected online journal usage statistics for the previous year is a significant variable in a potential model to predict future citations. Although the amount of variability (.05) explained by the model seems low it is relatively high for a single variable in a model analyzing a large set of count data. Overall, usage is an important variable in a potential multi-variable model for predicting citations.

The regression models also indicated that both publisher provided usage statistics and locally recorded usage statistics are significant variables in a model predicting citations. The high correlation between *Access Use* and *Publisher Use* indicates that the latter would be a valid substitute for the former when modeling citations for those institutions that do not or cannot measure use locally. In addition, further examination of this relationship has promise for validating usage statistics provided by publishers with usage statistics collected at the local level or for modeling missing publisher provided usage data with citations or locally collected usage statistics.

Higher citation counts of online journals indicate that online availability influences citation behavior. The current study holds promise for confirmation of this

behavior in the disciplines studied but more analysis is needed to confirm it in other disciplines. The provision of access and discovery tools for online journals did not have the same strong results. While some disciplines had journals that did significantly increase in citations after the availability of access tools, many others did not show significant results. The hypothesis that increased ease of access will influence citation behavior could not be confirmed for most disciplines, either because it is not yet apparent or does not actually exist. Repeating this analysis with more years of data may shed further light on this hypothesis, since a time-delay effect may exist for acceptance of such tools and the measurable appearance of increased citations. Furthermore, researchers may be able to find and use materials regardless of provision of access tools by librarians to a large degree, or at the very least, citation analysis does not reflect an increase in accessibility through more citation. The final tests run on publisher provided usage statistics lends strong support to the latter conclusion, for the journals in that sample clearly showed large and significant z-scores, indicating that access tools do affect usage measures.

Two caveats to the current study that may influence this type of analysis but could not be accounted for were the possibilities that the university's authors were citing more materials per published article or that overall article publication increases might influence local citation behavior. The publication patterns, even for an entire discipline, at a single university fluctuate wildly from year to year, and even when combined into larger two or three year groups, do not indicate significant changes in publications produced. At the university under study,

faculty publications (as measured by ISI) increased 15% from 1997 to 2003, but were almost entirely due to a large increase in publications from one discipline. Indeed, most disciplines produced fewer publications in the most recent three years than in the preceding three years. In addition, increases in the total number of articles provided may influence citations rates and partially explain higher citations in recent years. Total articles published in the journals included in the dataset did increase from 1997 to 2003 by 11%, but this change is not easily interpreted. Article publication peaked in 2003, but remained steady from 1999 to 2003, which encompasses most of this study's analyses. Total citations per article was another interesting aspect that emerged from an examination of the data. An increase in this metric, jumping from a constant 24-27 from 1997 through 2000, jumped to 30-33 from 2001 to 2003. This warrants further study and may indicate that provision of online journals and associated access tools influence citation behavior by allowing authors to review more materials per published article. In addition, the total number of online journal articles available locally was not available during the study time period. An overall increase or sharp increase of available online articles in some disciplines due to digitization of back issues, which was occurring as more and more publishers brought their journals online, may have affected the number of citations.

The analysis and study of usage measures is an important facet of collection analysis in today's academic research libraries. Escalating costs and renewed focus on providing high value resources with fewer financial resources at their disposal has librarians and administrators seeking new ways to select, deselect, promote, and justify research journal purchases. This analysis will help librarians to understand the relationship and predictive nature of usage measures collected locally and those supplied by information providers. It also draws together some aspects of the vast literature of citation analysis with the emerging research on online journal usage statistics.

The research shows that measurable use of resources is related to citation counts. This provides a foundation for developing models to cross-validate usage measures, model missing data, and perform similar analysis on other types of usage measures. In practice, librarians can use the techniques employed in this analysis to verify that measures supplied by information providers are reliable and valid in the absence of any ability to audit the collection and reporting of online journal usage by publishers, who may have a strong incentive to over-report usage. In addition, librarians can model missing data for the large number of journals that still do not have usage reported or for gaps in data collection methods.

Overall, this study can help librarians to differentiate between pure raw counts of journal usage and citations, put into context the numbers that we use for collection decisions, and provide a basis for collection of usage measures. Analysis like this could be incorporated into the development of statistical packages, collection assessment and analysis tools, and serials evaluation

tools that can assist librarians in real-time evaluation of their collections and provide information to faculty, administrators, and funding agencies that is otherwise not well understood.

**Acknowledgements:** I would like to gratefully thank Amelia McDonald, Michael Sedano, Phil Davis, David Mundo, and the faculty of the Department of Information Studies, UCLA for their diligent editing and helpful comments on early versions of this paper. I would also like to thank the two anonymous reviewers for their insightful suggestions.

## References

- Antelman, K. (2004). Do open-access articles have a greater research impact? *College & Research Libraries*, 65, 372-382.
- Black, S. (2005). Impact of full text on print journal use. *Library Resources & Technical Services*, 49, 19-26.
- Blecic, D. D. (1999). Measurement of journal use: an analysis of the correlations between three methods. *Bulletin of the Medical Library Association*, 87, 20-25.
- Bonzi, S. & Snyder, H. W. (1991). Motivations for citation: A comparison of self citation and citation to others. *Scientometrics*, 21, 245-253.
- Borgman, C.L. & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology*, 36, 3-72.
- Brooks, T. A. (1986). Evidence of complex citer motivations. *Journal of the American Society for Information Science*, 37, 34-36.
- Brown, C. M. (1999). Information seeking behavior of scientists in the electronic information age: astronomers, chemists, mathematicians, and physicists. *Journal of the American Society for Information Science*, 50, 929-943.
- Brown, C. M. (2001). The coming age of e-prints in the literature of physics. *Issues in Science and Technology Librarianship*. Available: <http://www.library.ucsb.edu/istl/01-summer/refereed.html>
- Cameron, A.C. & Trivedi, P.K. (1998). *Regression analysis of count data*. New York: Cambridge University Press.
- Cano, V. (1989). Citation behavior: Classification, utility, and location. *Journal of the American Society for Information Science*, 40, 284-290.
- Chrzastowski, T.E. & Olesko, B.M. (1997). Chemistry journal use and cost: results of a longitudinal study. *Library Resources & Technical Services*, 41, 101-111.
- Cole, J. R. & Cole, S. (1972). The Ortega hypothesis. *Science*, 178, 368-375.

- Darmoni, S. J. & Roussel, F. (2002). Reading factor: a new bibliometric criterion for managing digital libraries. *Journal of the Medical Library Association*, 90, 323-327.
- Davis, P. M. (2002a). Patterns in online journal usage: Challenging the composition of geographic consortia. *College & Research Libraries*, 63, 484-497.
- Davis, P. M. (2002b). Where to spend our e-journal money? Defining a university library's core collection through citation analysis. *portal: Libraries and the Academy*, 2, 155-166.
- Davis, P.M. (2004). Information seeking behavior of chemists: a transaction log analysis of referral URLs. *Journal of the American Society for Information Science*, 55, 326-332.
- De Groote, S. L. & Dorsch, J. (2003). Measuring use patterns of online journals and databases. *Journal of the Medical Library Association*, 91, 231-240.
- Feitelson, D. G. & Yovel, Y. (2004). Predictive ranking of computer scientists using CiteSeer data. *Journal of Documentation*, 60, 44-60.
- Gardner, W., Mulvey, E.P., & Shaw, E.C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological Bulletin*, 118, 392-404.
- Gilbert, G. N. (1977). Referencing as persuasion. *Social Studies of Science*, 7, 113-122.
- Goodman, D. (2002). A year without print at Princeton, and what we plan next. *Learned Publishing*, 15, 43-50.
- Harter, S.P. (1998). Scholarly communication and electronic journals: an impact study. *Journal of the American Society for Information Science*, 49, 507-516.
- Joswick, K. & Stierman, J. (1997). The core list mirage: A comparison of the journals frequently consulted by faculty and students. *College & Research Libraries*, 58, 48-55.
- Kaplan, N.R. & Nelson, M.L. (2000). Determining the publication impact of a digital library. *Journal of the American Society for Information Science*, 51, 324-329.
- Kaske, N.K. (1993). Research methodologies and transaction log analysis: issues, questions, and a proposed model. *Library Hi Tech*, 11, 79-86.
- Ke, H., Kwakkelaar, R., Tai, Y., & Chen, L. (2002). Exploring behavior of e-journal users in science and technology: transaction log analysis of Elsevier's ScienceDirect OnSite in Taiwan. *Library & Information Science Research*, 24, 265-291.
- Kelland, J.L. & Young, A.P. (1994). Citation as a form of library use. *Collection Management*, 9, 81-100.
- Langlois, D.C. & Von Schulz, J.V. (1973). Journal usage survey – method and application. *Special Libraries*, 64, 239-244.
- Luther, J. (2000). White paper on online journal usage statistics. Council on Library & Information Resources. Available: <http://www.clir.org/pubs/reports/pub94/contents.html>.
- MacRoberts, M. H. & MacRoberts, B. R. (1987). Another test of the normative theory of citing. *Journal of the American Society for Information Science*, 38, 305-306.
- MacRoberts, M. H. & MacRoberts, B. R. (1988). Author motivation for not citing influences: A methodological note. *Journal of the American Society for Information Science*, 39, 432-433.
- MacRoberts, M. H. & MacRoberts, B. R. (1989). Problems of citation analysis: A critical review. *Journal of the American Society for Information Science*, 40, 342-349.
- Mercer, L. S. (2000). Measuring the use and value of online journals and books. *Issues in Science and Technology Librarianship*. Available: <http://www.library.ucsb.edu/istl/00-winter/article1.html>.
- Merton, R. K. (1973). *The sociology of science*. Chicago: University of Chicago Press.
- Millson-Martula, C. (1988). Use studies and serials rationalization: a review. *Serials Librarian*, 15, 121-136.
- Montgomery, C. H. (2000). Measuring the impact of an online journal collection on library costs: a framework and preliminary observations. *D-Lib Magazine*, 6. Available: <http://www.dlib.org/dlib/october00/montgomery/10montgomery.html>.
- Morse, D. H. & Clintworth, W. A. (2000). Comparing patterns of print and online journal use in an academic health science library. *Issues in Science and Technology Librarianship*. Available: <http://www.library.ucsb.edu/istl/00-fall/refereed.html>
- NISO Standards Committee AY. Z39.7 Library Statistics Standard Revision Update: Library Statistics Data Dictionary. Available: [http://www.niso.org/committees/committee\\_ay.html](http://www.niso.org/committees/committee_ay.html)
- Nisonger, T.E. (1999). JASIS and library and information science journal rankings: A review and analysis of the last half-century. *Journal of the American Society for Information Science*, 50, 1004-1019.

Project Counter. (2002). Available:  
<http://www.projectcounter.org/>.

Rogers, S. A. (2001). Online journal usage at Ohio State University. *College & Research Libraries*, 62(1), 25-34.

Scales, P.A. (1976). Citation analyses as indicators of the use of serials: a comparison of ranked title lists produced by citation counting and from use data. *Journal of Documentation*, 32, 17-25.

Seglen, P. O. (1997a). Citations and journal impact factors: questionable indicators of research quality. *Allergy*, 52, 1050-1056.

Seglen, P. O. (1997b). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, 314, 498-502.

Seglen, P. O. (1998). Citation rates and journal impact factors are not suitable for evaluation of research. *Acta Orthopaedica Scandinavica*, 69, 224-229.

Shin, E. (2003). Do impact factors change with a change of medium? A comparison of impact factors when publication is by paper and through parallel publishing. *Journal of Information Science*, 29, 527-533.

Stemper, J.A. & Jaguszewski, J.M. (2003). Usage statistics for electronic journals: an analysis of local and vendor counts. *Collection Management*, 28, 3-22.

Tenopir, C. & King, D.W. (2002). Reading behaviour and online journals. *Learned Publishing*, 15, 259-265.

Tenopir, C. & King, D.W. (2004). Medical faculty's use of print and electronic journals: changes over time and in comparison with scientists. *Journal of the Medical Library Association*, 92, 233-241.

Tsay, M.Y. (1998). The relationship between journal use in a medical library and citation use. *Bulletin of the Medical Library Association*, 86, 31-39.

Vinkler, P. (1987). A quasi-quantitative citation model. *Scientometrics*, 12, 47-72.