User-defined valued metrics for electronic journals

Katherine Chew, James Stemper, Caroline Lilyard, Mary Schoenborn

University of Minnesota Libraries

Abstract

Purpose: Building on the work done by the California Digital Library (CDL), the University of Minnesota Libraries is developing a set of user-defined value-based electronic journal usage metrics. User value is assessed in three overall categories: (1) utility or reading value, (2) quality or citing value, and (3) cost effectiveness. In addition to analyzing vendor-generated usage metrics, also included were Affinity String data, derived from the University of Minnesota's central authentication system that anonymously captures a user's academic department and degree program or position at the university and combined with vendor-generated usage data, provides a granular picture of journal use down to the title level. Collection management librarians and library users can benefit from a viable, more accurate metric for use and value of library resources than cost-per-download, which would ensure that the most needed/valued resources are available to further research and learning.

Methodology: Metrics were identified that are utilized to determine e-journal retainability: OpenURL link resolver requests for article views, COUNTER-compliant downloads, JCR Impact Factors, Eigenfactor Scores, local citations from Thomson Reuters Local Journal Use Reports and Affinity String requests for article views. Two years of usage data were assessed using Pearson correlation coefficients to compare the different metrics. Affinity String data is correlated with the results to determine any discipline or degree level differences. A composite score is assigned to each journal to assess its overall value in comparison to other journals within the same broad subject category.

Findings: This project found SFX clickthroughs a more consistent predictor than COUNTER downloads of the journals our faculty will cite in their articles, with Eigenfactor a more consistent predictor of citation behavior than Impact Factor.

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

Introduction

For years, libraries have searched for the perfect usage metrics to help make tough decisions on what journals to retrain when budgets get tight. Download statistics do not tell the whole story; just because an article is downloaded does not mean it is later read or cited and cost per use or download does not always resonate with faculty. Available usage metrics all have their own unique limitations: OpenURL link resolvers may not capture all of the downloads, and COUNTER-compliant data is not available from all publishers or vendors. Value metrics such as Impact Factors are based on a short time interval that does not reflect the citation patterns of all disciplines, and they can be manipulated to some extent. The goal of our project is by comparing and combining a variety of usage metrics, the value that our users assign to our collection through their decisions about which journal articles to download, read and cite can be more accurately determined.

Background

The academic library community has grappled with metrics for e-journal management since the late 1990s. The literature reflects increasingly sophisticated attempts to capture data on usage. Homegrown, open-source, and commercial assessment systems exist to capture citation metrics, data collection usage, etc. A 1998 article in Serials Librarian, *Inferring user behavior from journal access figures*[1], outlined measures of e-journal usage. The "search to browse" ratio was analyzed, concluding SBR "provides a more meaningful measure than simple access count, but numbers alone will never provide a complete picture." *Evaluative usage-based metrics for selection of e-journals*, in College and Research Libraries[2] draws on existing work in the print world (Meyers and Fleming's "reasonably equitable quantitative evaluation tool"). The authors developed evaluative metrics: average cost per access, average cost per article, and content-adjusted usage. Coupled to benchmarks, a resource can be compared with a peer product, even when content and value differ. Studies from the early 2000s[3] demonstrate evolution of

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

standards for accessing e-content using tracking methods, while Lui and Cox[4] detail early attempts to organize standards to collect use data at the local level.

A Cornell University study in the Journal of American Society for Information Science and Technology[5] analyzed individual online behavior from individual IP addresses as a surrogate. Previous research came from citation analysis, library circulation studies, survey research and publisher statistics. An IP analysis of 29 online journals from the American Chemical Society revealed a majority of IPs target a small number of journals and downloads. Article downloads and users were strongly enough related to imply a user-population can be estimated by a journal's total use.

Another correlate is in the 2003 article "*Usage Statistics for Electronic Journals: An Analysis of Local and Vendor Counts*[6]. It compared locally-tracked title-level login statistics (pre-SFX) to download statistics from four major publishers. Local login counter's journal use rankings matched COUNTER statistics 70% of the time, suggesting local statistics are a viable alternative to vendor statistics.

An IFLA article[7] reviewed early attempts to develop cost per use data and discusses ways libraries, consortia, and publishers can use unit cost information for management decisions. Findings were from Drexel University, University of Muenster, University of Connecticut, and two major publishers (Emerald and IOPP). The advent of SFX, an OpenURL link resolver, in 2005 provided a reporting package for statistical data on successful and unsuccessful article view requests[8].

The MESUR project (Mellon Foundation) constitutes a systematic effort to define, validate and cross-validate a range of usage-based metrics of scholarly impact. "*Towards usage-based metrics: first results from the MESUR project*"[9], indicates MESUR has collected nearly 1 billion usage events and all associated bibliographic and citation data from significant publishers, aggregators and institutional consortia to construct a large-scale usage data reference set. An informal survey in May 2010 showed that standardized methods for e- journal data collection are emerging, but not yet cross platform/vendor[10]. Vendors are attempting to provide solutions[11].

Standardization is occurring. In March 2012, Project COUNTER released a draft Code of Practice for Usage Factors: Journals (UFJ). In this proposed standard, the UFJ will be "the Median Value of a set of ordered full-text article usage data for articles published in a journal during a two calendar year Publication Period"[12].

Wendt Library at UW-Madison surveyed engineering faculty for important criteria in journal cancellations. The journals they cited most in articles was ranked as most important, then journals they published in, usage statistics, impact factors, and affiliation with professional associations; the least important was cost per use, the metric many librarians use[13].

Assessment methods are becoming more sophisticated. California Digital Library (CDL) staff measure the value of journal titles through metrics in three categories: Utility (usage statistics and citations), Quality (Impact Factor and Source Normalized Impact per Paper (SNIP)), and Cost Effectiveness (cost per usage and cost per SNIP). A Weighted Value Algorithm assesses the journal's value in the institutional context, also factoring in disciplinary differences[14]. Our project builds on this by adding affinity string data to assess disciplinary differences.

The Data

The data for this project was collected from over 4,700 journals owned by or accessible to the University of Minnesota Libraries users. A recent CIBER research report[15] found that two years of journal usage data is sufficient to provide insight into a journal's usage patterns. The project dataset includes information for the two year period 2009 through 2010. The variables collected were: (1) online use, as tabulated by the library's OpenURL link resolver, SFX, (2) online use, as reported by publisher COUNTER-compliant reports, (3) University of Minnesota authors publishing activity, from Thomson Reuter's Local Journal Use Reports (LJUR), (4) Journal Citation Reports (JCR) Impact Factors, (5) Eigenfactor Scores and (6) locally-generated authentication logs known as affinity strings. Cost per use was calculated for each title for all the variables and included in the dataset.

It should be noted that data could not be collected for all of the variables for every title, as not every publisher is COUNTER-compliant, nor are there Impact Factors, Eigenfactors or LJUR data for every title due to Thomson Reuter's limited title indexing. Subject codes were attached to each title in order to do analysis within broad subject categories.

Some Definitions/Terminology

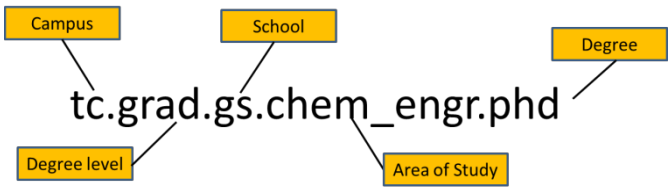
The variable, *SFX Downloads*, is the OpenURL link resolver count of "clickthrough" article view requests. SFX clickthroughs do not capture all of the local usage of the journals, as patrons can bypass the **FindIt** icon 

by using search engines like Google or bookmarking journal websites while they are searching or accessing resources within the campus' library IP ranges. A 2009 log analysis showed that 65% of library site use originated from off-campus[16]. An ongoing study by the University Libraries IT on library use and student success indicates between 47%-74% of undergraduates use online resources, while 65 to nearly 100% of the graduate level students utilize online resources [17]. Anecdotal and affinity string evidence puts faculty online resource use at a similar level as that of graduate students.

Vendor Download is the COUNTER-compliant publisher reporting of downloads of full-text articles. We used Journal Report 1, number of successful full-text article requests. [18] Publisher-supplied COUNTER reports capture more information about journal usage than SFX downloads, as publishers record activity via library-authorized IP addresses, including off-campus use via proxy server IP , search engines like Google Scholar and bookmarked sites. However, not every publisher is COUNTER-compliant, whereas, every University-licensed journal title does trigger SFX download counts. *Citation Counts* are the number of articles and citations by University authors in journals indexed by Thomson Reuter's Web of Science and were matched to titles subscribed to by the University of Minnesota Libraries. These counts come from a series of reports purchased from Thomson Reuters known as Local Journal Usage Reports (LJUR). The reports have three components: 1) Source counts per published year per journal title of articles authored by University faculty and staff; 2) Cited By, counts per publication

year per journal title where University authors cited others' journal articles in their own papers; and 3) Citing, counts per year of the citing article per journal title where other authors have cited our University authors' published works.

Affinity Strings are generated by the University's OIT using information from the University's human resources management system. All University students, staff and faculty are assigned affinity strings that are based on his or her area of work or study. An example of an affinity string would look like this: **ahc.fac.med**. Tracking this string backward, this represents the affinity string for a **medical** school **faculty** member within the **Academic Health Center** at the University. A more complex affinity string may look like this: **tc.grad.gs.chem_engr.phd**, which translates to a Twin Cities (tc) graduate school (grad /gs) PhD. student in Chemical Engineering.



Affinity strings provide a rich lode of information about user behavior, and aggregating similar affinity strings without personally identifying any one user does not compromise privacy. They allow for a granular picture of journal use down to the title level. Affinity string data is captured every time a person logs into the University central authentication system with their unique Internet ID to access library resources. As with SFX data, not all journal downloads are captured due to some users working within the campus' libraries IP ranges. However, this limitation has been changing, due to increased computer security measures across campus, which now directs that all campus computers require a login, including all staff, faculty and public computers.

Journal Citation Reports Impact Factor, from Thomson Reuters, is a measure of the frequency with which the average article in a journal has been cited in a given period of time relative to other journals in the field. The annual Impact Factor for any particular journal is a ratio between citations and recent

citable articles and is calculated based on a three-year period. It is the average number of times published papers in a journal are cited up to two years after publication[19]. Though highly discipline-dependent and subject to manipulation, as noted, it does provide a sense of journal citation patterns.

Eigenfactor Score, developed by Jevin West and Carl Bergstrom at the University of Washington [20], is an attempt to calculate the likelihood that a journal is going to be used, or how much time a user will spend accessing content from that journal. The Eigenfactor Score is based on the number of times published articles are cited from a journal within the past five years of article citations, as oppose to the three years of Impact Factors. Journal self-citations are removed and weight is given to articles from more prestigious journals. The Eigenfactor algorithm tries to account for differences in influence among citing journals and also adjusts for differences in citation patterns among disciplines.

The Questions

Utilizing the CDL's work on developing a set of user-defined value-based metrics, our project framed the following questions:

1) Utility or "reading" value: Does SFX clickthrough data combined with affinity string data provide a "good enough" departmental view of user activities, such that COUNTER-compliant data is expendable?
2) Quality or citing value: which is the better metric for representing value through citation behavior by users – JCR Impact Factor or Eigenfactor? Do either SFX clickthroughs or COUNTER downloads sufficiently correlate with local citation patterns via Local Journal Usage Reports (LJUR) or external citations via impact factor metrics? If both of them do, are impact factors expendable?
3) Cost effectiveness or cost value: how should these reading and citing values be combined with cost data to create a "cost-per-activity" metric that meaningfully informs collection management decisions?

Methodology

We began our analysis by combining all of our variables into a single spreadsheet. The journal title list, with publisher imprints, format (online/print), ISSNs, subject fund code and 2010 invoiced cost was downloaded from the University Libraries' ILS and its serials agent EBSCO into an Excel spreadsheet. The title list was then matched by ISSNs to SFX "clickthroughs", COUNTER-compliant downloads, LJUR data, JCR Impact Factors, Eigenfactor Scores and affinity string data and added to the spreadsheet. We cleaned up our data as much as possible: identifying titles that were missing SFX, COUNTER or affinity string download data due to mismatches on ISSNs; missing cost data due to publisher bundle pricing or where the electronic cost was located on the print record; duplicate records were removed, as well as any title that did not have electronic access in some form, such as print only or microform formats.

As this project's questions are framed around comparing variables and their relationships, correlation analysis was the method chosen to examine these relationships, similar to a study conducted in 2006 that compared print and online journal usage[21].

Averages for the 2009-2010 data years were calculated for the variables *SFX and COUNTER downloads*, *Affinity Strings*, *Impact Factor*, *Eigenfactor*, and the *LJUR* reports, to allow for variable comparisons. Data analysis was done using R[22] a well-known and freely available integrated statistical software suite for data manipulation, calculation and graphical display.

The spreadsheet was converted into an R data file and zero and blank values on the spreadsheet that represented unavailable data (where a zero in the COUNTER data field meant "N/A" as opposed to zero downloads) were compensated for with algorithms.

Scatter plots were created to determine if there were any positive, negative or no correlational relationships between variables and what the significance of these relationships was. We looked specifically at Pearson's correlation coefficient (signified by r), the coefficient of determination (r^2) and

p-value. *R* reflects the degree of linear relationship between variables, ranging between +1 to -1. A correlation of +1 indicates that there is a perfect positive linear relationship, whereas -1 indicates a perfect negative relationship and zero implies no relationship. Though there can be a range of interpretation depending on the discipline, it is generally accepted within the social sciences that an $r < 0.35$ is considered a low to weak correlation, 0.36-0.67 modest or moderate, 0.68-1.0 strong or high correlations, with anything over 0.90 a very high correlation (Table 1).[23]

Table 1: Range of Pearson Values

Correlation	Negative	Positive
None	-0.09 to 0.00	0.0 to 0.09
Low	-0.3 to -0.1	0.1 to 0.3
Moderate	-0.5 to -0.3	0.3 to 0.5
Strong	-1.0 to -0.5	0.5 to 1.0

In conjunction with the correlation coefficient, we looked at the coefficient of determination, which is the square of *r* and is reported as r^2. R^2 is the proportion or percent of variation in the values of the variables, or the degree of linear association between two variables. R^2 is often expressed as a percentage. So if an $r^2=0.34$, it would mean that 34% of the variance of *x* to *y* is explained, but also that 66% is unexplained[24].

P-value is the estimated probability of rejecting the null hypothesis, where the smaller the *p-value*, the more strongly is the rejection of the hypothesis of "no difference", or that the variables have no relationship whatsoever. *P-values* of less than 0.05 can be considered statistically significant, with $p<0.001$ statistically highly significant, or less than one in a thousand chance of being wrong[25].

Results and Discussion

To answer our question on utility or "reading" value, we began by running Pearson correlations between SFX clickthroughs, COUNTER downloads and affinity strings year by year and then as an average of the two years combined. Our sample size for all of our the correlations was n=4782, unless otherwise noted.

We found that for the SFX/COUNTER average correlation r=0.72, r^2=0.52 and p-*value*<2.2e-16 (fig. 1).

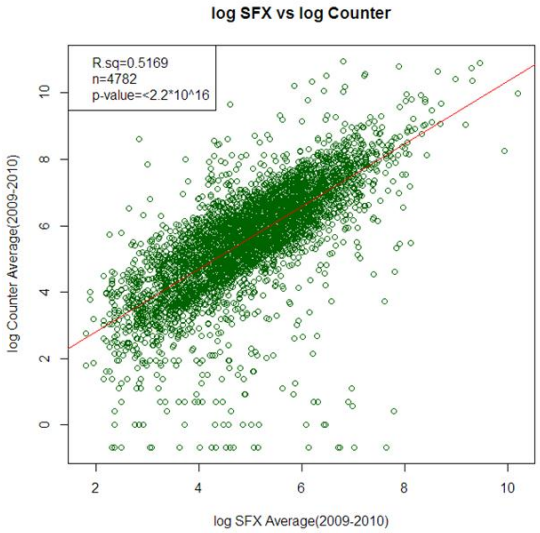


Figure 1: Average SFX and COUNTER

All of our correlations' F test *p-values* were less than 2.2e-16 (2 x 10^{-16}). Since our *p-values* are very small, this indicates that the model we chose is "statistically highly significant" and the variables are linearly related to each other. Based on our correlation table, the r^2 value and its accompanying scatter plot show a strong, positive correlation between SFX and COUNTER. Specifically, 52% of the variance in the COUNTER article downloads from each journal can be explained by the SFX clickthrough requests for each journal. Conversely, 48% of the variance in the COUNTER

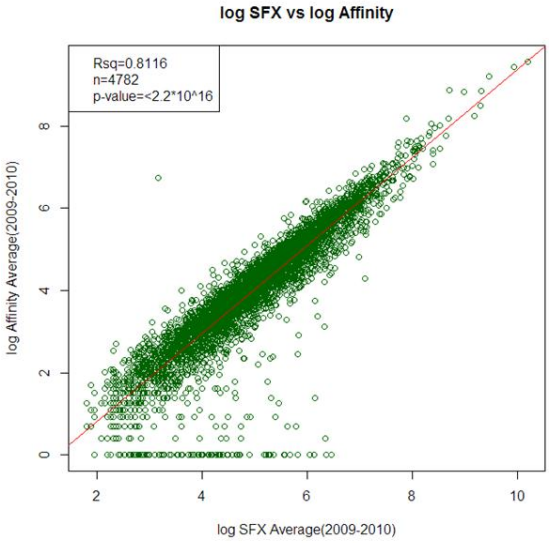


Figure 3: SFX and Affinity Strings

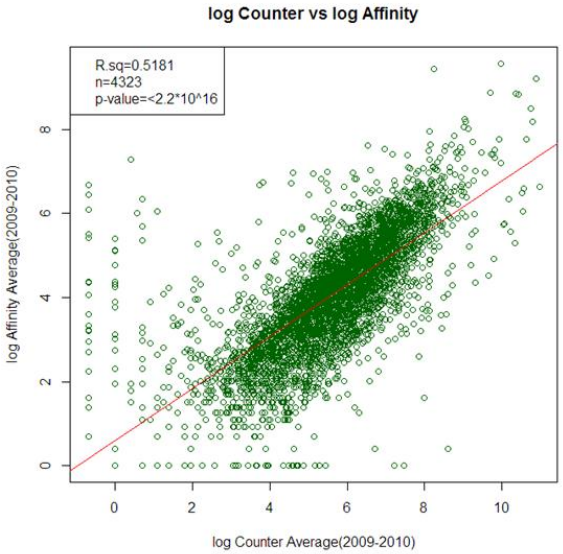


Figure 2: COUNTER and Affinity Strings

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

article downloads can only be explained by other factors. SFX explains little more than half of the variance between the two ranked lists of most/least downloaded journals. When an additional correlation was run to account for potential COUNTER data errors, such as lower or non-existent downloads for titles with SFX clickthroughs, the r-squared rose to an astounding 74%.

The relationship between affinity strings and SFX is very strong, as would be expected, with r=0.90 and r^2=0.81 (fig.2).

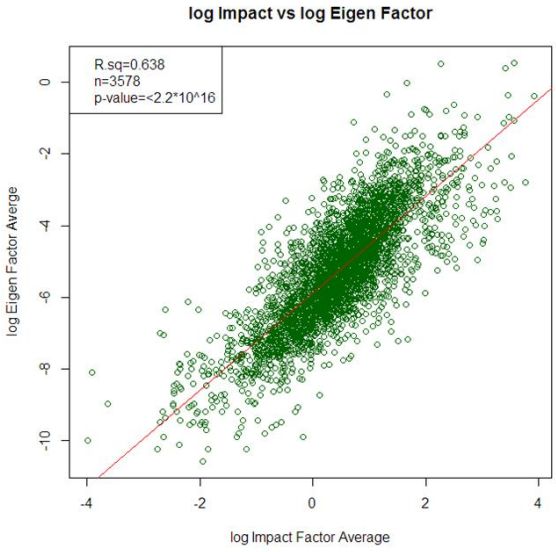


Figure 4: Impact Factor and Eigenfactor Scores

The relationship between affinity strings and COUNTER, though strong according to our correlation table, is slightly more than two-thirds of SFX and COUNTER where r=0.72, and r^2=0.52 (fig.3).In answering question number two on quality or citing value, first we ran correlations between Impact Factors and Eigenfactor Scores to see which was the better metric for citation behavior, and then compared SFX and COUNTER with U of M author citation behavior patterns via the Local Journal Usage Reports (LJUR). The correlation between the average of the 2009-2010 Eigenfactor Scores and Impact Factors for our title list is r=0.64, (n=3579),and r^2 = 0.80 (fig. 4). Like the CDL, we found that Eigenfactor compares very strongly to Impact Factor at 80%, with only 20% of the variance explained by other factors. Eigenfactor data is publicly available on the web; it only costs the staff time to download it and integrate it into your dataset.

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

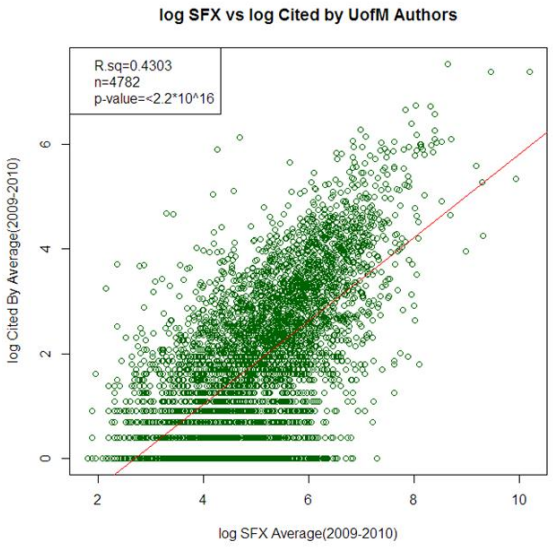


Figure 5: SFX and Cited By U of M Authors

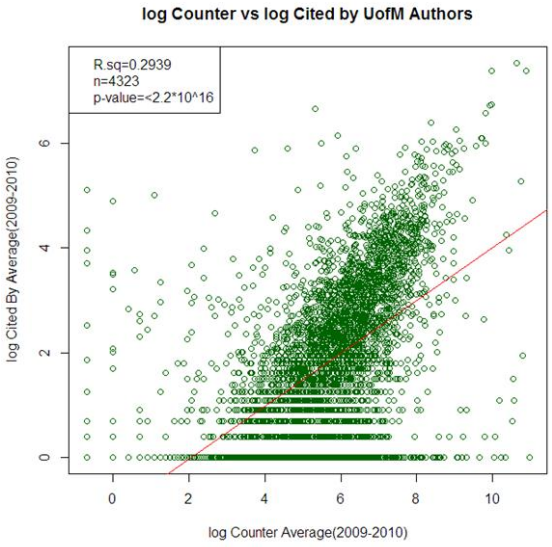


Figure 6: COUNTER and Cited By U of M Authors

The correlation between the average of the 2009-2010 SFX clickthroughs and Cited By section of the LJUR is r=0.66, and $r^2 = 0.43$ (fig. 5). The corresponding correlation between the average of the 2009-2010 COUNTER downloads and the Cited By LJUR is r=0.54, and $r^2 = 0.29$ (fig. 6). SFX is a moderate predictor of the journals our faculty will cite in their articles. COUNTER explains only 29% of the variance between the two ranked lists of most/least accessed and cited journals, and SFX explains at least

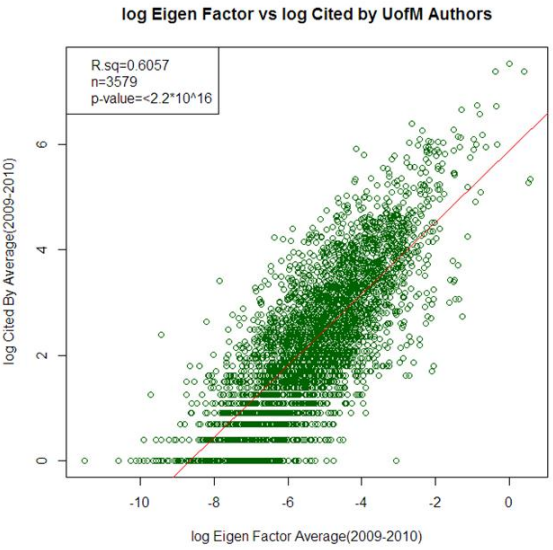


Figure 7: Eigenfactor Scores and Cited By U of M Authors

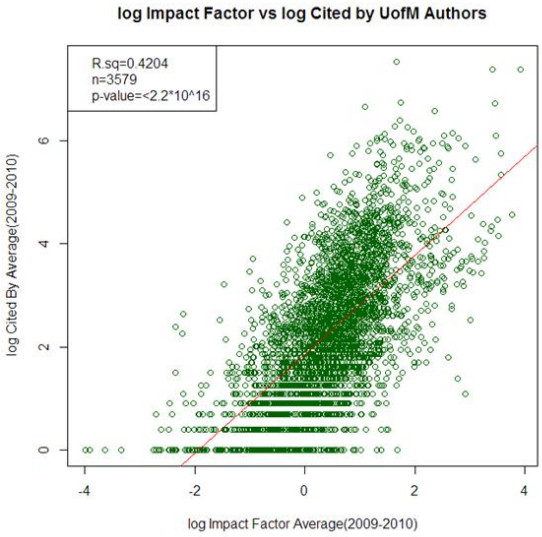


Figure 8: Impact Factor and Cited By U of M Authors

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

43% of the variance, leaving more than half of the variance (57% with SFX, and 71% with COUNTER) to be explained by other factors.

The correlation between the average of the 2009-2010 Eigenfactor Scores and Cited By LJUR statistics is r=0.81, (n=3579), and r^2 = 0.65 (fig.7). The correlation between the average of the 2009-2010 Impact Factors and Cited By LJUR statistics is r=0.65, and r^2 = 0.42 (fig.8).

Impact Factor explains 42% of the variance between the two ranked lists of most/least cited journals, and Eigenfactor explains 65% of the variance. Eigenfactor explains the majority of the variance, while Impact Factor explains a moderate variance. Eigenfactor is a more consistent predictor than Impact Factor of the journals our faculty will cite in their articles -- about a third more reliable. Notably, both metrics are better predictors of our faculty's citation behavior than our traditional download statistics. So it is worth going beyond our traditional "cost per use" (i.e. "cost per download") calculations when deciding which journals to keep and which to cancel.

Disciplinary Case-Studies: Humphrey School of Public Affairs , Marketing Department of the Carlson School of Management, Academic Health Center School of Nursing

In these case studies we tried to answer the question as to what degree faculty, students and staff associated with a particular program or school – as opposed to other academic departments – are using journals acquired on their behalf from funds specifically designated for that discipline/school or program.

Three schools/departments were chosen to test our questions, two from the social sciences and one from the health sciences, to see if there are any disciplinary differences in journal usage behavior.

The School of Nursing has 57 full time faculty and a student headcount of nearly 900, with about 400 undergraduates and 500 graduate students. The program includes a Bachelors, two Master's programs,

and two doctorate programs. In the student success project, nearly 100% of the nursing students and faculty used online resources.

The Humphrey School of Public Affairs offers five master degree programs, six dual degree programs, as well as executive certification programs covering a broad range of policy and planning areas. The school has approximately 28 full time faculty and a student headcount of 500. The public affairs journals are comprised of a targeted collection of core titles intended to support foundational policy and planning research needs.

The Carlson School of Management includes Bachelors, Masters and doctorate programs. Total student enrollment is approximately 4,500. The Marketing Department is one program of the Carlson School and has approximately 15 full time faculty and a degree-seeking student headcount of approximately 335 students.

For each department/school, a list of journals in the corresponding subject fund, along with SFX, COUNTER, affinity strings , Impact Factors, Eigenfactor and Cited By LJUR were pulled from the master list. Correlations were run between the two utility/reading values, SFX and COUNTER and our locally available affinity strings; then between each of the three utility/reading values (SFX, COUNTER, Affinity Strings) and Cited By; and finally between each of the two quality/citing values (Impact Factor, Eigenfactor) and Cited By (Table 2).

Table 2: R-Squared Values for Disciplinary Case Studies

R^2 Values	Humphrey School of Public Affairs	School of Nursing	Marketing Department
SFX/Affinity Strings	0.6426	0.8101	0.796
COUNTER/Affinity Strings	0.3606	0.4786	0.825
SFX/Cited By	0.3571	0.2464	0.6914
COUNTER/Cited By	0.50323	0.3914	0.7614
Affinity String/Cited By	0.5248	0.1375	0.9076
Impact Factor/Cited By	0.1691	0.3666	0.3981
Eigenfactor Score/Cited By	0.4796	0.6363	0.8723

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

For all of the case studies, correlation between SFX and affinity strings represented strong relationships (fig 9). The Humphrey School of Public Affairs correlation was 65%, the Carlson School of Management Marketing Department was very strong at 80%, and the School of Nursing was also very strong at 81%. SFX's very strong 81% correlation within Nursing affinities is almost 20% higher than the Public Affairs and only slightly higher than the Marketing Department affinities. Meanwhile, the affinities' correlation with COUNTER downloads shows the Marketing Department with a very strong 83%, followed by Nursing with a more moderate 48% and a low of 36% for public affairs (fig.10). This data suggest that SFX is most certainly "good enough" when combined with affinity strings, and thus the unavailability of COUNTER data for some publishers is not critical.

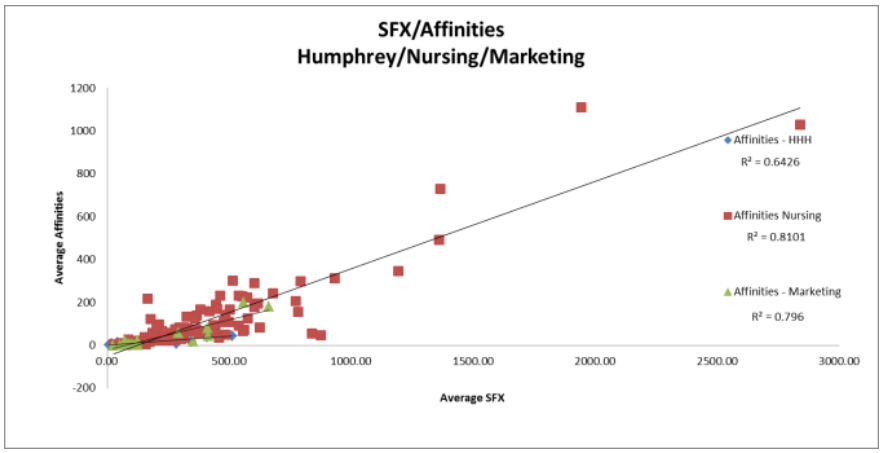


Figure 9: SFX and Disciplinary Affinity Strings

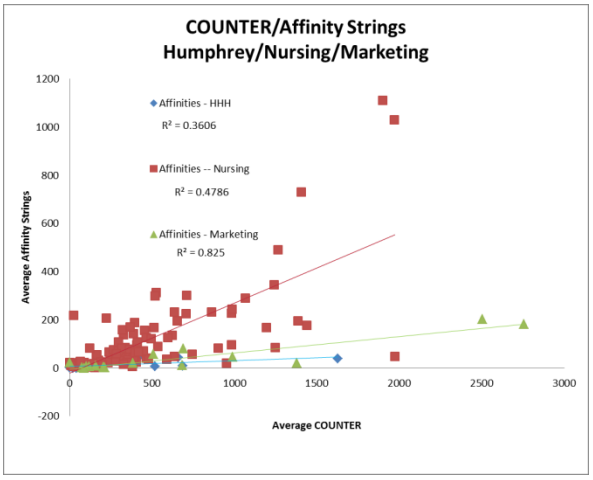


Figure 10: COUNTER and Disciplinary Affinity Strings

When looking at the reading value metrics, both nursing and public affairs journals have a low correlation of SFX downloads with local faculty citation behavior at 25% and 36%, respectively.

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

Marketing journals on the other hand, shows a strong 69% (fig 11)

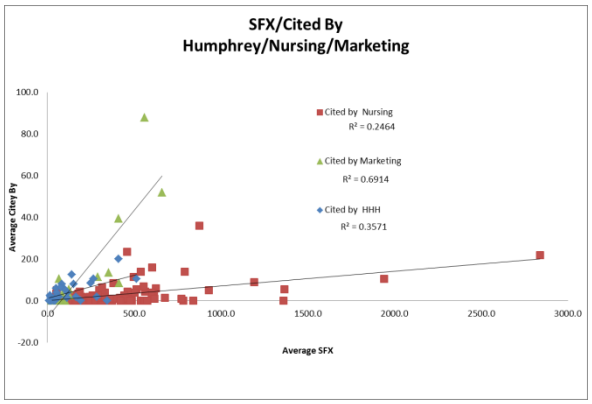


Figure 11: SFX and Disciplinary Cited By

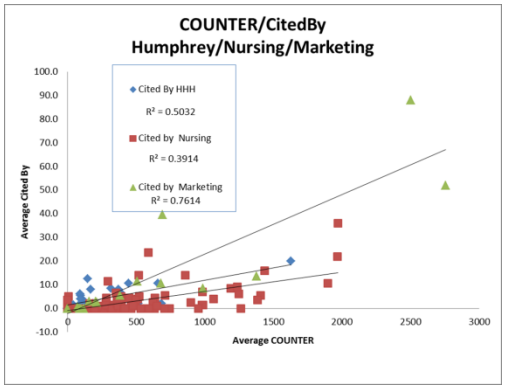


Figure 12: COUNTER and Disciplinary Cited By

Despite the better marketing journal performance correlation with SFX, the data clearly shows even better correlations for all three case subject collections with COUNTER downloads where they range from a high of 76% for marketing journals to a very moderate 39% for nursing journals (fig. 12). So, despite SFX correlating better than COUNTER with faculty citation behavior in the overall journal set, the opposite was the case with these three disciplines, indicating that a selector's decision dataset should include COUNTER if it is available from the subject's publishers. The correlations of the departmental affinity data with LJUR show Marketing once again with the best predictive relationship, at a very strong 91%, followed by Public Affairs at 52% and Nursing at a very low 14%. These results indicate that while an overall analysis can show one pattern, individual disciplines can vary markedly in which reading measures best correlate with faculty citation behavior. The variance in ISI's coverage by discipline suggests itself as a reason, though we also found that eliminating aggregator journals from the subject list before running departmental correlation had the effect of strengthening the correlation. Turning next to citing value metrics, a very strong and consistently significant correlation for Eigenfactor over Impact Factor is evident across our three case study subjects. The marketing journals are very strong at 87%, with nursing at a moderately strong 64%, and public affairs journals at a moderate 48% (figs 13 & 14).

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

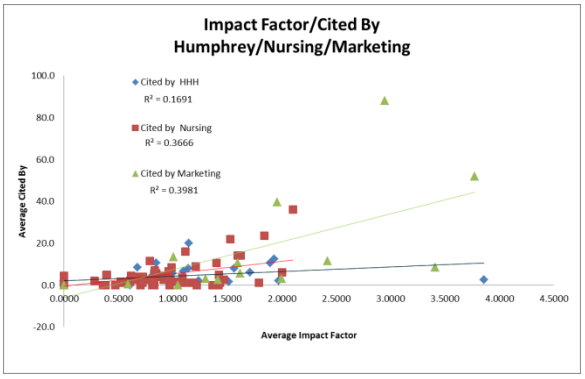


Figure 13: Impact Factor and Disciplinary Cited By

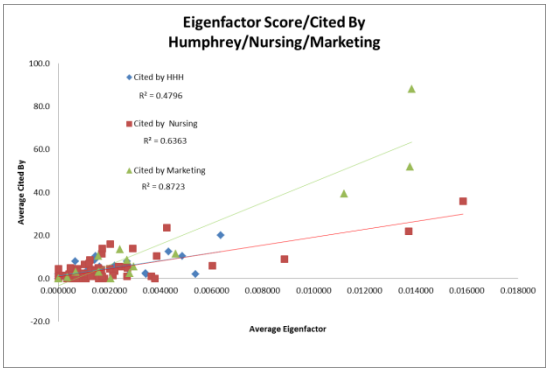


Figure 14: Eigenfactor Scores and Disciplinary Cited By

Yet, even at this modest level for public affairs journals, the Eigenfactor score is significantly more predictive than Impact Factor. These subject level citation use results align with our overall journal set analysis and findings.

When looking at subsets of the correlations as a functional group, patterns emerge. Marketing journals had, by far, the strongest correlations for both the reading value metrics and the citing value metrics; Public Affairs the second strongest correlations in reading values, and Nursing the weakest correlations in that area. However, they switched places in the citing value metrics, with Nursing having the second strongest correlations in citing values, and Public Affairs the weakest correlations in that area. So the weighting that each selector gives to a type of metric -- the degree to which s/he has confidence it will gauge the potential citability of a journal -- will vary by subject.

One can rank the combined five reading and citing metrics for each subject as to the strength with which they correlate with local faculty citation. The affinity strings were the *most* predictive of citations in the two social science case studies. COUNTER downloads were the *second most* predictive of citations in 2 of the 3 case studies (social science and a hard science subject). SFX was the fourth most predictive of citations for all of the case studies, contradicting the findings of the overall journal set, where it fared much better, while Impact Factor was the least predictive of citations in 2 of the 3 case studies (the social science subjects).

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

With just these three case studies, we have a preliminary answer to our question about whether or not there may be discipline differences in journal usage. The data from these case studies demonstrate evidence of variation not only when each is compared with the overall journal set, but also shows marked differences in journal usage patterns between the selected case subjects. Yet, despite the correlation strength variation between disciplines, there exists a remarkable consistency at the subject level, whereby all show that COUNTER downloads are a better predictor of citation behavior than SFX requests. Similarly, the cases are consistent with the overall journal set findings of the strength of Eigenfactor score. These case data provide for more granular analysis (the trees vs. the forest) that can inform collection decision making at the discipline level.

Conclusion

The foundational question of this project was to determine which readily available journal usage data from various sources provides evidence of value to our academic users. Value was assessed in three categories: (1) utility or reading value, (2) quality or citing value, and (3) cost effectiveness.

Our initial findings from statistical analysis of local journal use indicate that SFX clickthrough data when combined with affinity string data does provide a "good enough" result. SFX is useful in that it provides evidence that electronic journals are being accessed, but it also does tell us much more. Correlating this information with affinity string data provides a stronger data set that illustrates a reasonable relationship between the journals purchased on behalf of an academic department and use by its intended audience.

However, the data does demonstrate that the inclusion of COUNTER data strengthens the positive correlation when looking at citation behavior. We found that COUNTER downloads are a more consistent predictor than SFX clickthroughs of the journals our faculty will cite in their articles, though not by much, between 10-25% depending on the discipline. Thus, obtaining COUNTER-complaint download data in addition to SFX clickthroughs provides a reasonable measure of local "reading value" ,

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

supports the strongest information about the journals our faculty use in their own research papers and a stronger collection management position.

With regards to our second question, which is the better metric for representing value through citation behavior by users – JCR Impact Factor or Eigenfactor Scores, our analysis indicates a very favorable relationship between the two. Eigenfactor Scores proved to be better predictors of University faculty citation behavior, by a large margin, with moderate to very strong correlations.

While Eigenfactor is free and Impact Factor remains a standard –bearer of journal significance, the results suggest stronger collection development decisions result when both data points are considered.

We included some discipline snapshot analysis to determine if patterns seen at the larger scale offer meaningful results at the disciplinary level. To achieve this result, the use-log data that is collected must identify "who" is using the materials, not just that the materials are being used. This additional data is critical to gathering relevant results.

The examples presented of a public affairs graduate school, nursing school and marketing program shows that there are disciplinary differences in SFX clickthroughs alone for subject funded titles, where the correlations range from 64% to a very strong 81%. This offers evidence that a reasonable relationship exists between the journals purchased on an academic journal fund and downloading by the intended audience. Citation behavior also shows a very marked disciplinary difference, where the correlations with affinity strings run from a very low 14% (nursing), to a very high 91% (marketing). On the other hand, when citation behavior is correlated with COUNTER downloads, there is an interesting flip in nursing journals, which rises from a very low 14% to a moderate 39%, whereas both public affairs and marketing slightly drop. This most likely is a sign of interdisciplinary use of nursing titles, which is another data point to consider when making collection management decisions.

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]

And finally, these data demonstrate that taking" reading" and citing values together provide better information than the traditional "cost per use" metric many have used that simply looks at what is being downloaded.

This project offers further evidence of impact when measures of journal use, broadly defined as incorporating both "reading" and citing, are compared with disciplinary ties of those actually doing the reading and citing at the local level. Selectors need to go beyond the convenience of quick and easy OpenURL resolver data divided by subscription price calculations if they are to justify increasingly difficult cancellation decisions to faculty and administrators who demand accountability.

Here is our proposed formula:

((Cost ÷ SFX) + (Cost ÷ Cites)) / 2

It incorporates both a reading metric and a citing metric. The more cites a title gets, the lower its cost per use. For a reading metric, we chose SFX because our overall results show that it is a "good enough" substitute for COUNTER; COUNTER can then act as a tie breaker if SFX totals for two titles are the same. For citing metric we chose LJUR, for practical reasons that it is well known and regarded, and it is customized to our institution; since our results show that Eigenfactor is an adequate substitute for Impact factor, Eigenfactor can also be a tie breaker if needed. This formula works best for those titles that are indexed in Web of Science (and thus in Eigenfactor as well) and for which we can thereby get reports on local citation behavior.

Next steps

We have amassed an abundance of data and have barely touched its potential. Some areas we would like to explore further are more discipline related case studies and delving more deeply into what the affinity strings can tell us about user behavior, such as faculty use versus graduate student or undergraduate usage patterns. Another would be vendor analysis -- which publisher's journals are cited more than others by our University authors? And lastly, going beyond the data available from the Web of Science indexing

and looking to Scopus data and SNIP, or Source Normalized Impact Per Paper, for additional usage metrics.

Acknowledgements

We would like to acknowledge the efforts of our data-wrangler, George Swan, the input of our library colleagues Jan Fransen, Shane Nackerud, and Amy West, and the sage advice and help of the University of Minnesota Statistical Consulting Clinic's Professor Sanford Weisberg and graduate statistics student Grace Jackson and the University of Wisconsin-Eau Claire Mathematics Department's Assistant Professor Abra Brisbin.

Endnotes

1. Alan Dawson, "Inferring user behavior from journal access figures." *The Serials Librarian* 35 no. 3 (1998): 31-41, DOI**:**10.1300/J123v35n03_03.

2. Karla L. Hahn and Lila A. Faulkner, "Evaluative usage-based metrics for selection of e-journals." *College & Research Libraries* 63 no. 3 (2002): 215-227.

3. Julie M. Hurd, Deborah D. Blecic and Anne E. Robinson, "Performance measures for electronic journals: a user-centered approach." *Science & Technology Libraries* 20 no. 2/3 (2001): 57-71.

4. Weiling Liu, Fannie M. Cox, "Tracking the use of e-journals: a technique collaboratively developed by the Cataloging Department and the Office of Libraries Technology at the University of Louisville", *OCLC Systems & Services*, 18 no. 1 (2002):32 – 39.

5. Philip M. Davis and Leah R. Solla, "An IP-level analysis statistics for electronic journals in chemistry: making inferences about user behavior." *Journal of the American Society for Information Science and Technology* 54 no. 11 (2003):1062-1068, DOI: 10.1002/asi.10302.

6. James A. Stemper and Janice M. Jaguszewski. "Usage statistics for electronic journals: an analysis of local and vendor counts." *Collection Management* 28 no. 4 (2003):3-22.

7. Brinley Franklin, "Managing the electronic collection with cost per use data." *IFLA Journal* 31 no. 3 (2005):241-248.

8. Tina E Chrzastowski, Michael Norman and Sarah Elizabeth Miller, "SFX statistical reports: a primer for collection assessment librarians." *Collection Management* 34 no. 4 (2009):286-303.

9. Johan Bollen, Herbert Van de Sompel and Marko A. Rodriquez, "Towards usage-based impact metrics: first results from the mesur project." In*: Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries*, *Pittsburg*, *PA*, June 16-20, 2008: 231-240, doi:10.1145/1378889.1378928.

10. Rachel A. Fleming-May and Jill E. Grogg, "Practitioner responses on the collection and use of usage statistics." *Library Technology Reports* 46 no. 6 (2010):28-34.

11. Ibid. 11-16.

12. The COUNTER Code of Practice for Usage Factors: Draft Release 1. (March 2012): 6. http://www.projectcounter.org/documents/Draft_UF_R1.pdf

13. Helman, Deborah. "Wendt Library – Faculty Collections Survey 2008." University of Wisconsin-Madison, 2008. Unpublished survey. Used by permission of the author.

14. Jacqueline Wilson and Chan Li, "Calculating scholarly journal value through objective metrics." *CDLINFO News*, February 13, 2012, http://www.cdlib.org/cdlinfo/2012/02/13/calculating-scholarly-journal-value-through-objective-metrics

15. CIBER Research Limited, *The Journal Usage Factor: exploratory data analysis: Stage 2 Final Report*, 27 May, 3 Aug 2012 http://www.projectcounter.org/documents/CIBER_final_report_July.pdf

16. Cody Hansen, Heather Hessel, John Barneson, Deborah Boudewyns, Jan Fransen, Lara Friedman-Shedlov, Martha Hardy, Chris Rose, Barb Stelmasik and Stacie Traill, " Discoverability Phase 1: Final Report." http://conservancy.umn.edu/handle/48258

17. Shane Nackerud, Jan Fransen, Kate Peterson, Kristen Mastel, Krista Soria and David Peterson, "Library data and student success." *Library Technology Conference*, Macalester College, St. Paul, MN, March 14-15, 2012. http://digitalcommons.macalester.edu/libtech_conf/2012/sessions/28/

18. Project COUNTER http://www.projectcounter.org

19. The Thomson Reuters Impact Factor. http://thomsonreuters.com/products_services/science/free/essays/impact_factor/

20. Jevin D. West, Theodore C. Bergstrom and Carl Bergstrom, "The Eigenfactor™ metrics: a network approach to assessing scholarly journals." *College and Research Libraries* 7 no. 3 (2010):236-244. http:// www.eigenfactor.org/

21. R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

22. John D. McDonald, "Understanding journal usage: a statistical analysis of citation and use." *Journal of the American Society for Information Science and Technology* 58 no. 1(2007):39-50.

23. Richard Taylor, "Interpretation of the correlation coefficient: a basic review." *Journal of Diagnostic Medical Sonography* 6 no. 1 (1990): 35-39.

24. James Dean Brown, "The coefficient of determination." *SHIKEN: JALT Testing & Evaluation SIG Newsletter* 7 no 1(2003):14-16.

25. Anthony C. Redmon, Anne-Maree Keenan, "Understanding statistics: putting p-values into perspective." *Journal of the American Podiatric Medical Association* 92 no. 5(2002):297-305.

Preprint of the following paper:
Chew, Katherine, Stemper, James, Lilyard, Caroline, Schoenborn, Mary. 2012. User-defined valued metrics for electronic journals. In: Proceedings of the 2012 Library Assessment Conference, Charlottesville, VA, Oct 29-31. [In press]