

Comparative Analysis of Citation Studies, Swept Use, and ISI's Impact Factors As Tools for Journal Deselection

Joseph G. Wible
Falconer Biology Library
Stanford University
Stanford, California

ABSTRACT

The serials collection at Stanford's Falconer Biology Library was analyzed to identify \$20,000 worth of titles to cut in 1989. Three methods were used to collect data. First, a citation analysis was done using all papers published over a two-year period by the library's primary clientele, the faculty of the Department of Biological Sciences. Second, swept use data for materials used in the library were measured during one-week sample periods during each of the fall, winter, and spring academic quarters. Finally, the impact factor for each title was taken from the Institute for Scientific Information (ISI) annual bibliometric journal analysis.

The collection of swept use and citation data is labor intensive compared to looking up ISI impact factors. However, relying solely on ISI's data would result in serious deselection errors. Among the 100 most heavily cited titles, 16 received an impact factor of less than one, and three were not even included in ISI's ranking. The citation and swept use data complement each other, and together provide a more complete picture of an individual title's value to the collection. For example, among the 100 most heavily cited titles, seven had no swept use during the sample period. Looking at the reciprocal data, titles with as many as 34 in-house uses during three weeks were not cited by the faculty for two years.

INTRODUCTION

Over the past several years, subscription rates to scientific serials have increased well beyond the rate of inflation (Ivins, 1988; Okerson, 1986; Holden, 1987; Houbeck, 1987) (Figure 1). As a result, most libraries are no longer able to accurately predict or financially support the book budget necessary to maintain existing buying patterns. During 1989 it became necessary for Stanford to cut science serial subscriptions by \$100,000. Of this amount, \$20,000 in cuts were to be made by Falconer Biology Library. At the time, the biology library at Stanford had a materials budget of \$326,881, of which an estimated 75% was used to maintain approximately 1311 active serial subscriptions.

In order to reach its \$20,000 goal, Falconer Biology Library cancelled subscriptions to 93 serial titles starting in 1989. Criteria for deselection included the journal's price, its availability elsewhere on campus or regionally, and the importance of the journal to the community being served. While the first two criteria are objective factors, the importance of a journal is subjective, less easy to quantify, and subject to interpretation. Three different measurements were used to assess the value of a journal to the collection: a site-specific citation analysis, a swept use study, and ISI's impact factor. The multiple approaches were used to determine whether the three measures were closely correlated and therefore provided redundant information. In particular, was ISI's impact factor a sufficiently valid measure of the importance of a journal to Falconer Library's collection, thus eliminating the need to perform labor intensive use and citation studies?

METHODS

Citation analysis consisted of identifying the number of times an individual journal title was cited in publications of the Biology Department's faculty over a period of two academic years (1984/85-1985/86). Data was collected manually, using a compiled set of the faculty's publications supplied to the library annually by the department. Using a printed list of the library's serial holdings, staff went through the bibliographies of the faculty papers, marking those titles cited. Also compiled was a list and count of serial titles cited by the faculty but not owned by the library. It might have been possible to automatically compile some of the data using the ISI database, but this would have incurred costs. Also, the faculty publish in titles not indexed by ISI. Since this work could be done while staff was on duty at the circulation desk, the manual approach was used.

Swept use was a measure of the number of times a bound journal was removed from the shelf during week-long sample periods during the 1980/81 and 1988/89 academic years. Sample weeks were distributed across each of the fall, winter, and spring quarters for a total of six weeks. Since Falconer Library does not allow its journals to circulate, this procedure provided a quantifiable measure of in-house use.

The **ISI impact factor** for each title was taken from the 1986 edition of ISI's bibliometric journal analysis. All data were entered into a serials database maintained on an IBM-AT using the relational DBMS program DataEase.

RESULTS

During the two-year sample period, 331 out of 1311 titles were cited at least once by the faculty (Figure 2). *Proceedings of the National Academy of Sciences* was the most heavily cited title at 527 times. During the swept use sample weeks, 479 out of 1311 titles were used at least once (Figure 3). With 448 uses, *Proceedings of the National Academy of Sciences* again topped the list (Figure 4).

A total of 98 serial titles for possible cancellation were presented to the faculty. Accompanying each title was its annual subscription price, the number of times it had been cited by the faculty during the study period, the number of times the faculty published in it during the study period, the number of times it was used in-house during the six week sampling, and a list of libraries where the title could be obtained (Figure 5). The faculty requested only five titles be dropped from the cut list. Of the 93 serials cancelled, ten were reference serials which were not included in the study. Of the remaining 83 titles, 48 had not been cited or used in-house during the course of this study. Only seven of the titles cancelled had been cited by the faculty, and all of these were duplicated in other Stanford libraries. Most of the 26 titles with at least one in-house use were also either duplicated on campus or available from local libraries with cooperative agreements with Stanford.

The citation analysis also identified 167 titles which were cited by the department faculty but were not owned by Falconer Library. Only one new subscription was initiated based on this list; most of the titles were cited only once and were either found in other libraries at Stanford or were clearly out of scope as defined by the library's collection development policy.

DISCUSSION

The citation and swept use data complement each other, and together provide a more complete picture of an individual title's value to the collection. For example, among the 100 most heavily cited titles, seven had no swept use during the sample period. It is known that most faculty maintain personal subscriptions to titles most relevant to their work, and these personal copies are often shared among the faculty's lab group. This practice provides one possible explanation as to the lack of correlation between the swept use and citation data. Looking at the reciprocal data, titles with as many as 34 in-house uses were not cited even once by the faculty. Remembering that the citation analysis numbers were derived only from biology department faculty papers, this difference could reflect the use of titles by non-biology department faculty or by undergraduate students, graduate students, information brokers and other visitors. One would also expect the faculty to be reading material beyond those which they later cite and to rely on the library's subscription, especially for the more expensive titles.

Figure 1
Sample Serial Prices 1987/1988*

TITLE	1987 PRICE	1988 PRICE	PERCENT INCREASE	1987 INFLATION RATE
<i>Biochemica et Biophysica Acta</i>	\$3,573.00	\$4,530.50	27%	4.40%
<i>Brain Research</i>	\$3,861.00	\$4,859.00	26%	4.40%
<i>Cell</i>	\$ 235.00	\$ 250.00	6%	4.40%
<i>Journal of Biological Chemistry</i>	\$ 455.00	\$ 480.00	5%	4.40%
<i>Journal of Molecular Biology</i>	\$1,250.00	\$1,368.00	9%	4.40%
<i>Nature</i>	\$ 250.00	\$ 275.00	10%	4.40%
<i>Proc. of the National Academy of Sciences</i>	\$ 215.00	\$ 250.00	16%	4.40%

*For further information concerning journal rates, see the articles listed in the bibliography.

Figure 2
Citation of Serial Titles
Frequency Histogram

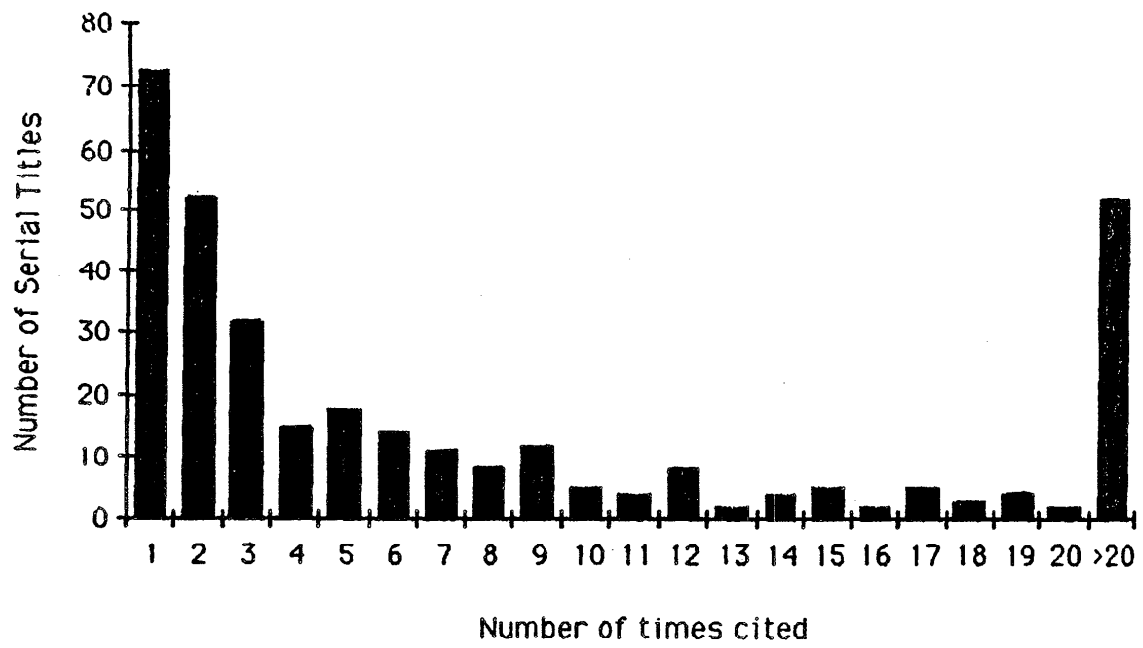


Figure 3
Swept Use of Serial Titles
Frequency Histogram

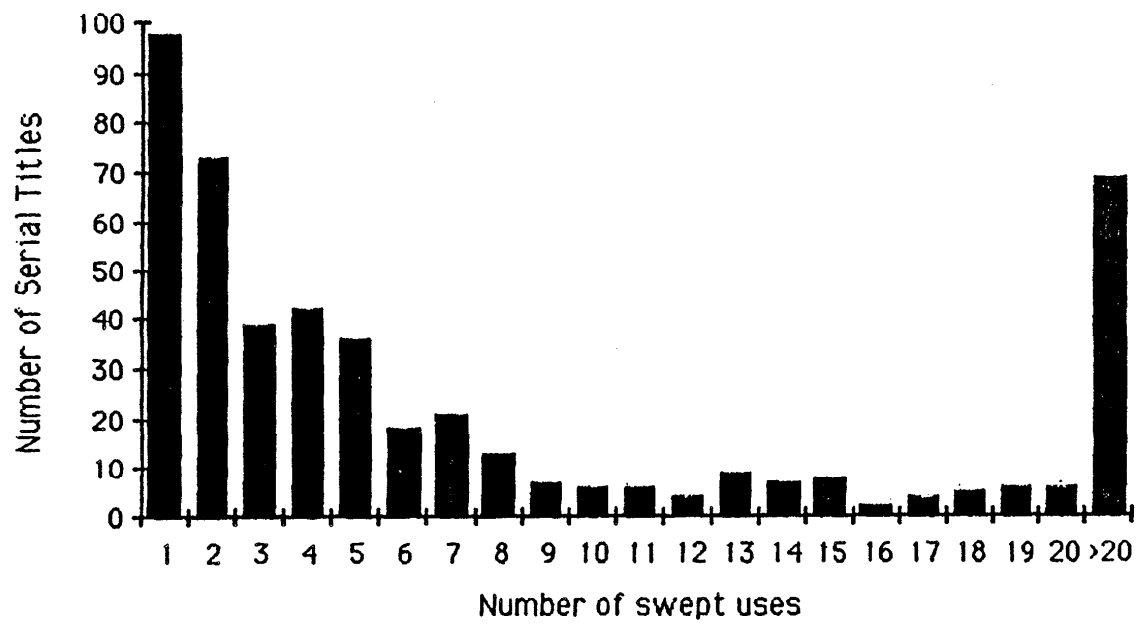


Figure 4

**Falconer Biology Library
Serial Data Sorted by Cite***

Price (\$)	Cite	Pub	Swept	Impact	Title
250.00	527	22	448	9.166	<i>Proc. National Academy of Sciences</i>
250.00	321	7	405	15.252	<i>Nature</i>
455.00	281	6	290	6.315	<i>Journal of Biological Chemistry</i>
250.00	228	4	200	20.098	<i>Cell</i>
195.00	208	4	89	3.054	<i>Genetics</i>
98.00	206	3	369	12.437	<i>Science</i>
1,297.66	178	11	47	1.727	<i>Oecologia</i>
	171	10	164	3.005	<i>Plant Physiology</i>
340.00	171	7	139	3.071	<i>Journal of Bacteriology</i>
	126	8	102	3.629	<i>Developmental Biology</i>
1,411.97	123	1	127	6.597	<i>Journal of Molecular Biology</i>
100.00	123	4	155	2.633	<i>Ecology</i>
90.00	122	4	81	2.878	<i>American Naturalist</i>
4,530.50	122	1	215	2.739	<i>Biochimica et Biophysica Acta</i>
	106	4	24	1.603	<i>Theoretical Population Biology</i>
575.98	104	2	118	6.055	<i>Nucleic Acids Research</i>
280.00	96	5	35	6.625	<i>Molecular and Cellular Biology</i>
1,252.11	94	2	56	2.982	<i>Molecular and General Genetics</i>
999.00	93	7	54	2.869	<i>Planta</i>
350.00	84	6	11	0.133	<i>Journal of Neuroscience</i>
100.00	69	1	47	2.798	<i>Evolution</i>
439.00	66	2	132	3.829	<i>Biochemistry</i>
40.00	62	0	34	5.658	<i>Ecological Monographs</i>
275.00	61	2	125	8.802	<i>Journal of Cell Biology</i>
	57	0	6		<i>Carnegie Institution of Wash. Yearbook</i>
2,346.00	55	1	71	1.988	<i>Mutation Research</i>
	55	0	42	1.653	<i>Methods in Enzymology</i>
130.00	51	1	46	2.765	<i>Cold Spring Harbor Symp. Quantitative Bio.</i>

*As of November 2, 1988

Cited = Times cited by department faculty during 1984/85 - 1985/86 academic years

Pub = Times faculty published in title during 1984/85 - 1985/86 academic years

Swept = Times used in library during three weeks in 1980/81 and three weeks in 1988/89

Impact = ISI's Impact Factor for 1986

Figure 5
Falconer Biology Library
Selected Serial Cuts for 1989

Price (\$)	Cite	Pub	Swept	Location Duplicated*	Title
728.00	0	0	0	Medical	<i>Vision Research</i>
768.00	0	0	5	Engineering	<i>Water, Air and Soil Pollution</i>
800.00	0	0	1	UC Davis	<i>Human Physiology</i>
1,298.00	13	0	9	Hopkins	<i>Marine Biology</i>
428.00	0	0	0	Hopkins	<i>Estuarine Coastal and Shelf Sciences</i>
60.00	1	0	1	Math	<i>Biometrics</i>
250.00	0	0	0	Medical	<i>Computers in Biology and Medicine</i>
405.00	0	0	0	Medical	<i>Cell Motility and the Cytoskeleton</i>
1,150.00	0	0	6	Medical	<i>Biochemical</i>
435.00	0	0	0	Chemistry	<i>International Journal of Biochemistry</i>
65.00	0	0	0	Physics	<i>Advances in Optical & Electron Microscopy</i>
330.00	0	0	0	UC Berkeley	<i>Environmental Biology of Fishes</i>
16.00	0	0	0	Gov. Docs.	<i>Fishery Bulletin (US Nat. Mar. Fish. Serv.)</i>

*All are Stanford libraries except University of California (UC) Berkeley and Davis campuses.

The compilation of citation and swept use data is labor intensive compared to looking up ISI impact factors. If the ISI impact factor could be used as the sole measure of the importance of a title to a library's collection, it would save the library time and money when faced with the task of collection evaluation. However, in the case of this library's collection, relying solely on ISI's data would result in serious deselection errors. For example, among the 100 titles most heavily cited by the faculty, 16 received an impact factor of less than one and three were not even included in the ISI rankings. It is possible that a larger, more comprehensive library collection may find a better correlation between the ISI impact factor and the value of a serial title to its collection. In the case of Falconer Library, the ISI impact factor was not used to justify journal deselection to the faculty.

The use of citation analysis to justify the maintenance of a serial subscription seems justified; if the title is being cited heavily by the library's primary clientele, it should be held in the library's collection even if it is not specifically the library's copy which is being used. This is a distinctly different use of citation analysis from one where it is used as a measure of the importance of an individual scientist's body of work. The latter is plagued with unproven assumptions (MacRoberts & MacRoberts, 1986). And while citation and swept use data are useful tools for identifying potential cancellations, one must exercise judgement. For example, a low citation and swept use total could merely reflect the fact that a journal is a relatively new title or a recent title change. Therefore, it is important to emphasize that evaluations as described here should not be used blindly, and that in the end, each title must be considered individually.

BIBLIOGRAPHY

- Holden, Constance. 1987. "Libraries Stunned by Journal Price Increases." *Science* 236:908-909.
- Houbeck, Jr., Robert L. 1987. "If Present Trends Continue: Responding to Journal Price Increases." *Journal of Academic Librarianship* 13:214-220.
- Ivins, October. 1988. "Serial Prices." *Serials Review* 14(3):61-66.
- MacRoberts, Michael H. and Barbara R. MacRoberts. 1986. "Quantitative Measures of Communication in Science: A Study of the Formal Level." *Social Studies of Science* 16:151-172.
- Okerson, Ann. 1986. "Periodical Prices: A History and Discussion." *Advances in Serials Management* 1:101-134.