

# Practical Interpretation of Citation and Library Use Studies

*Most studies of journal citations and library use are of little if any practical use to librarians and information system designers because of inadequate data collection and analysis. The paper considers the data required to guide (a) the librarian in acquisition (current and retrospective), discarding, and binding; and (b) the information system designer in selecting journals to be scanned for secondary services, selecting items from journals scanned, and retiring items from active files.*

**T**HIS PAPER IS CONCERNED only with the units in which citation and use analyses need to be expressed if they are to be of practical value. Questions such as the validity of citations as indicators of use, size of sample needed (of sources and of actual citations) in view of the wide variability between journals and volumes within journals, the relative usefulness of different sources of citations, the differences between synchronous and diachronous data, and the problems of forecasting future use, have been discussed elsewhere.<sup>1,2</sup>

Citation and library use studies have most commonly considered journals, counting the frequencies of use or citation, calculating "obsolescence" rates, ranking titles in order of citations received, and so on. These studies are claimed to be of use to librarians and information system designers, in guiding them as to what to buy and when to discard. As these studies have been re-

ported, however, most are of limited value, and some are positively misleading. The data have too often been collected with inadequate sampling techniques, forgetting that an age-group or title with many volumes is more likely to appear in a random sample in which volumes or papers are being counted than one with fewer. Ranked lists of crude "uses" are valueless; most do not even take into account the length of time each journal has been in existence. Before useful conclusions can be drawn, results must be expressed in units that allow for distorting factors, e.g., not as frequencies of use, but as densities of use per item, and so on.

To assess the practical value of use studies, it is desirable first to consider what decisions a librarian or information system designer may need to make that could be aided by a use study.

## LIBRARY PROBLEMS

To take questions faced by the librarian first:

1. Which journals to buy. The librarian will have a strictly limited budget or will want to know which 500 (or 1,000 or 2,000) journals

---

*Maurice B. Line is director general, British Library Lending Division, and Alexander Sandison is assistant keeper, British Library, Science Reference Library.*

will give him best value for money. To assess this he needs to know *how many uses per monetary unit* each journal provides. The costs of a journal include the initial ordering cost and then annually recurring costs of:

- a. Subscription;
- b. Accounting, claiming parts not received, etc.;
- c. Receiving and processing of parts—this depends on the number of parts per year;
- d. Preparing for binding, and actual binding—depending on the bulk received each year (number of volumes and size of volumes);
- e. Storage cost—depending on the additional shelf space occupied each year, which for many titles fluctuates substantially, and for some was increasing exponentially during the 1960s.

So far, to the best of our knowledge, no studies giving this information have appeared. Reanalysis of data from a citation study, to take account of the above factors, can produce large differences in rank order.<sup>3</sup> It should also be noted that some studies have shown substantial changes in rank from year to year;<sup>4</sup> these changes are especially likely to affect the middle ranking journals, for which the selection decision is in any case the most difficult.

2. What volumes of which journals to discard, and when. For both his space problems and retention costs, the librarian needs to know when *the number of uses per unit of shelf space* is at a level where retention is uneconomic (though he may still decide to retain for other reasons). There seems to be so much variation between individual journals in the use made from year

to year that generalized "obsolescence rates" are of no value whatever. Data have, therefore, to be collected for individual journals. When this is being done, the uses of each volume have to be related to the space it occupies—if a 1970 volume receives twice as many uses as a 1950 volume, and it is twice as thick, both are earning their keep equally. Such statements as "the half life of physics literature is four years," so far from being of practical use, are extremely dangerous, since they take no account of several important factors. A recent reanalysis of data from a library use study of physics journals showed that once allowance was made for the steadily increasing bulk of many journals, there was no sign of "obsolescence" except in the very recent issues.<sup>5</sup> There were also substantial differences between titles in the level of use per unit of shelf occupation.

In certain cases, library space may be so strictly limited that the librarian will be more concerned with uses relative to other material than with any arbitrary level of actual use. If this is so, he may wish or need to dispose of material that is strictly earning its keep, but the data required for comparing one type of material with another still need to be expressed as uses per unit of shelf, or uses per dollar, or a similar appropriate unit.

3. Whether or not to bind. This question is really an extension of the previous two. Some journals are heavily used for "updating" purposes for some months after they are first received and only rarely for "basic" or retrospective searches after they are more than two or three years old; they may not be worth binding at all. Any library

use study, however, which counts unbound parts and bound volumes as if they were equal units is useless here, for they must inevitably inflate the apparent relative use of the most recent issues.

4. Whether to buy a back run. For this, an assessment has to be made of the relative costs and benefits of borrowing and buying. This depends on the date beyond which it is less economic to buy than to borrow, on the shelfspace required for the back run, and the purchase price. None of this information is provided by citation studies, or by use studies conducted in other libraries, and the purchase price in particular may vary a good deal from supplier to supplier. The only value of citation studies is that they may be of some small aid in providing evidence on the need for older volumes, but they will probably underestimate the relative use of the older material. Interlibrary loan demands from the library in question are likely to provide better evidence.

#### INFORMATION SYSTEM PROBLEMS

The information system designer is faced with three problems: selecting the material to scan; selecting items to make available in indexes, abstracts, or computerized current awareness services; and selecting, retaining, and relegating items in manual or machine-readable files for retrospective searching.

1. Which selecting journals should be scanned for articles to include. Here the cost factors are very different from those faced by the librarian. The service may not in fact purchase the journals, so that the concern is with numbers of items extracted per unit of indexing time, which is in its turn influenced by the use of titles or key-

words in the system, the explicitness of the titles, and whether the summaries are sufficiently informative for the services' special interest. There is likely to be such a close link between an item's coverage in indexing services and its citation by authors, that analyses of citations can be of very little use for comparing journals already indexed with those not yet scanned. In any case, all such analyses need to be as densities of citations per citable item (Garfield's impact factor<sup>6</sup>). Comparisons of such densities for the same titles before and after inclusion in an indexing service would be of great interest.

2. Which items should be selected for current awareness. In this case, the newness of items precludes the existence of citation or library use data. But surveys of the occurrence of keywords per search profile might perhaps be relevant.
3. Which records to "retire" from an active file, and when. The system designer needs some measure of the number of relevant recalls per unit of memory store or of the number of times particular keywords appear in search profiles. In judging whether a recall rate is too low for retention in an active file, the great variations in the fashionableness and topicality of search topics must be remembered, and there should be provision for recovering relegated records from the "passive" file as interests change. Nearly all systems have been designed on the tacit assumption that items a few years old are of little interest. The few library use studies that have related use data to the amount of material available for use suggest that this may be a fallacy and that, for retrospective searching, greater age

does not necessarily indicate lower relevance or value.

#### CONCLUSION

As will have been seen, very few citation or library use studies so far have given helpful information. This is partly because the data need to be collected locally on uses of the local library or file. However, published studies could go much farther than they do, first, in pointing out their severe limitations and, secondly, in providing such information as uses per subscription cost, uses per foot of shelf, uses per article, recalls per keyword, etc. What would be of great value is a collection of information on the size and growth of a wide range of individual journals in terms of articles, of pages, and of thickness (unbound, but without advertisements, etc.).

#### REFERENCES

1. J. Michael Brittain and Maurice B. Line, "Sources of Citations and References for Analysis Purposes: A Comparative Assessment," *Journal of Documentation* 29:72-80 (March 1973).
2. Maurice B. Line and Alexander Sandison, "'Obsolescence' and Changes in the Use of Literature with Time," *Journal of Documentation* 30:283-350 (Sept. 1974).
3. Maurice B. Line, "Optimization of Library Expenditure on Biochemical Journals," *Journal of Documentation* 31:36-37 (March 1975).
4. Alexander Sandison, "The Use of Older Literature and its Obsolescence," *Journal of Documentation* 27:184-99 (Sept. 1971).
5. Alexander Sandison, "Densities of Use, and Absence of Obsolescence in Physics Journals at MIT," *Journal of the American Society for Information Science* 25:172-82 (May-June 1974).
6. Eugene Garfield, "Citation Analysis as a Tool in Journal Evaluation," *Science* 178:471-79 (Nov. 1972).