

Using Indexes as "Memory Assists"

A number of studies have demonstrated that readers at all levels of library proficiency use the library catalog and other reference tools primarily to track down books and articles that they have learned about elsewhere—from book reviews, bibliographies, recommended reading lists, and friends and colleagues. Their tendency is to look things up by author or title, whichever is better known or remembered. The comparative usefulness of several well-known reference tools for this purpose is the subject of this report.

THE TWO RESEARCH PROJECTS reported here investigate the usefulness of index services when they function as "memory assists."

The studies are part of an argument that goes like this: Very useful services can be offered to the university and research community by providing computer access to the "hard" elements in a bibliographic reference. These elements are title words, author, journal, year, sponsoring agency, etc.—elements that can be extracted from an article or report by relatively unskilled clerical help. Subject indexing is a highly skilled art, and the construction and maintenance of thesauri are expensive undertakings.

In fact, a very good case has been made by H. F. Dammers and others for de-emphasizing the intellectually demanding task of providing such specialist indexing for input to data files and to rely more on easily obtained and more prosaic computer products.¹

Several studies have shown that university faculty and graduate students use the reference collection (indexes, abstracts, etc.) of a university library to look up items already known to them either through their own reading or from some other source. They tend to use the reference collection to a lesser degree for general browsing or subject searching.² If this is so, providing memory assists makes sense.

TITLE WORD SEARCHING IN TWO WILSON INDEXES

Many libraries subscribe to the indexes published by the H. W. Wilson Company, presumably because they can function in both ways—as memory assists and as subject search tools. This study attempts to discover whether the Wilson *Social Sciences Index* and the Wilson *Humanities Index* (herein referred to as one index, *SSHI*) can be used to conduct searches for specific journal articles when title, or approximate title, is known but neither author or journal can be readily called to mind.

In a real-life situation, this type of search is conducted by someone wishing to verify a bibliographic citation or to retrieve an article for rereading when the author is not remembered. Such a search is easy to make in index services that provide title word indexing, such as the *Social Science Citation Index (SSCI)*. However, *SSCI* is a relatively new service and is not available everywhere. The question to be answered is: Even though neither the *Humanities Index* nor the *Social Sciences Index* formally provides title search capability, can they be used as an acceptably efficient substitute for a computer-produced title word index?

Methodology

Five areas in the social sciences were selected for study: sociology, anthropology, history, political science, and economics. Through a table of random numbers, three faculty members were selected in each of these disciplines from the lists published in the Temple University catalog. In four cases

Elaine C. Clever is head, circulation department, Temple University Library, Philadelphia, Pennsylvania.

repeats were necessary because the original person selected was not available for interview. In this way, fifteen professors were obtained who were willing to assist in the study.

Interview schedules were arranged during which the professors were asked to name two journals they read regularly and to select a recent (1974 or 1975) article from each that they remembered reading and that they considered to be of some importance to their work.

In nine cases the journals were available in the professor's office, and the interview took place there. In six instances the interview was conducted in the university library because the journals were regularly scanned there. Thus thirty articles for searching were selected from twenty-five journals (there were five duplicate choices).

After making a choice of two articles, the professors were asked to supply up to three subject terms under which they would probably look up each article if they were searching for it. They were informed that the searches would be made in *SSHI* and that if they were familiar with that index, they were at liberty to tailor their terms to it. In all, fifty-three terms were created to be used in the search for the thirty journal articles. This list is called list A.

In addition to the list A index terms, two to three words to be used as search terms were selected from each title. This second list, designated list B, was derived in a "mechanical" way by an undergraduate student who was instructed to underline the significant words in each title. List B contained fifty-one terms.

These two lists, developed in contrasting ways, were used as the instruments to test the ability of *SSHI* to retrieve the original articles. Since the lists were developed independently, they could also be expected to yield some information on the respective merits of each type of indexing, intellectual or mechanical.

Results

Journal Coverage. The journal coverage of *SSHI* was very good. Only four of the twenty-five journals in this selection were not regularly indexed by *SSHI*. However, six of the original articles appeared in those

four journals, leaving only twenty-four articles out of thirty that had any possibility of turning up in the search. Since it was assumed that neither author nor journal information was available to the searcher, the original thirty articles must be retained in calculating user effort.

However, the number of articles that could have been retrieved was not the theoretical twenty-four, but nineteen. An author search was conducted to see what the real possibility was of retrieving the articles that appeared in the twenty-one journals covered by *SSHI*. That five articles were not picked up for indexing is probably accounted for by the time lag between the date of publication of a missing article and its appearance in the index.

Intellectual versus Mechanical Indexing. The next question is: How many articles were actually retrieved by list A and list B?

The intellectually derived search terms in list A retrieved a total of nine articles. This is a user effort ratio of 30 percent (nine out of thirty). It is a recall ratio of 47 percent (nine out of nineteen). Since these index terms were supplied by active practitioners in each of the five disciplines included here, there seems to be some discrepancy between the skill level of *SSHI* and that of these particular experts.

How well did searching by mechanically derived index terms do in *SSHI*? The answer, somewhat surprising, is slightly better, but not to a significant degree. The fifty-one title words retrieved thirteen articles. The comparisons are shown in table 1.

TABLE 1

COMPARATIVE PERFORMANCE
OF "INTELLECTUAL" AND
"MECHANICAL" SEARCH TERMS

List	Hits/30	Hits/19
List A (Intellectual)	30%	47%
List B (Mechanical)	43%	68%

Using SSHI as a Memory Assist. Assuming that there had been a 100 percent correspondence between the search terms used by the researcher (whether the terms were derived intellectually or mechanically) and the subject headings in *SSHI*, the results would have been considerably better. Nine-

teen of the original thirty articles were indeed indexed by *SSHI*. But, of course, this is an unrealistic assumption. Long familiarity with *SSHI* subject headings and terms might have improved the showing for the effort expended on list A but would have made no difference for list B.

As far as this study goes, the results seem to show that *SSHI* is designed more for straight subject searching than for assisting in the retrieval of specific items. However, even here it is interesting to note that searching remembered title words may yield results that are somewhat better than searching by topic.

Note: The interview technique employed in this project was extremely time consuming, despite the fact that the number of indexers was small (fifteen). The original intention had been to recruit indexers from eight social science fields, but time constraints ruled this out. It should be noted also that perhaps this was not a large enough professorial sample to ensure that their indexing abilities were fairly representative of the faculty as a whole.

AUTHOR SEARCHES IN *SSHI* AND *SSCI*

Although it is true, as Allen Kent has argued,³ that scholars are finding it virtually impossible to read all of the pertinent literature to which they should be exposed if they are to keep abreast of developments in their fields, and that therefore they must increasingly rely on highly sophisticated tools and subject search strategies to keep informed, it is also true that habits are hard to change and that many scholars continue to use index services for author searches.⁴

This study is concerned with the usefulness of representative index services for author searches as they might be conducted by research scholars in the social sciences and humanities. It is one aspect of a study to investigate data bases in these fields and, as such, is partial in nature.

Two widely held and widely respected services were selected for study. They are the Wilson *Social Sciences Index* and *Humanities Index* (*SSHI*), once again considered as a single service, and the computer-produced *ISI Source Index* volume of the *Social Science Citation Index* (*SSCI*). The latter index is a newcomer to

the field, having been introduced to the public in 1973, and is a sister to the *ISI Science Citation Index*.

Methodology

Carl M. White's *Sources of Information in the Social Sciences*⁵ was used as the source to obtain a representative sample of journals from which to select look-up articles. White divides the social sciences into nine topics or disciplines: social science literature, history, geography, economics and business administration, sociology, anthropology, psychology, education, and political science.

Under each of these headings there is a section entitled "Sources of Scholarly Contributions" that lists the primary journals in each field, a total of 294 journals. These are presumed to be the ones most likely to be seen by scholars actively engaged in research and teaching.

A table of random numbers was used to select two to three journals from each of the pages in the sections. (The sampling method yielded an average representation of 18 percent of the journals listed in each of the nine categories. The range was from 10 percent in geography to 22 percent in history.) The resulting list was then used to select, once again at random, a recent issue from the current periodicals shelves of the Temple University library. The original list contained forty-seven journals. All but two were available. The remaining forty-five journals were considered to be a large enough sample to conduct the tests so that no effort was made to obtain the two missing journals from another library.

In the next step of the procedure one article or report was selected from each of the forty-five journal issues by application of a table of random numbers to the table of contents. The resulting authored references became the basic list to be searched. In addition, each of the forty-five journals was also scanned to secure one or two recent references (1973 or later) from footnotes or bibliographies. In some cases, notably in the history journals, no recent reference could be found. Therefore, the second list to be searched contains fewer items. Five journals appear in both lists, but this second set also contains journals that are neither

duplicated in the first nor mentioned in White.

However, since the items represented in this set presumably could have been the object of an author search by the reader of the article in which they are referenced, or by a scholar who read the article at the time of its original publication, they are included in the set. In the discussion that follows, the first set, containing forty-five items, will be referred to as list I, and the second set, containing thirty-eight items, will be referred to as list II.

Results

The first question to be answered is how well each of the services covers the journal literature as represented by list I and list II. The answer is shown in table 2.

TABLE 2

COVERAGE OF JOURNALS BY SSCI AND SSHI

	SSCI	SSHI
List I (45 items)	42 (93%)	33 (73%)
List II (38 items)	30 (78%)	18 (47%)

In list I, forty-two out of forty-five journals, or 93 percent, are regularly scanned by SSCI while thirty-three, or 73 percent, are covered by SSHI. List II, which contains a number of less prestigious journals, does not fare as well. SSCI picks up thirty out of thirty-eight, while the score for SSHI is only eighteen out of thirty-eight. Thus the possibility of turning up an author from either list is much higher in SSCI than in SSHI.

This result is to be expected considering the number of journals each service covers. SSCI lists 1,272 journals, while the combined *Social Sciences* and *Humanities* indexes cover 523 journals. All other things being equal, including such economic factors as subscription rates, SSCI would be a clear winner if journal coverage alone were the critical factor.

The possibility of finding a bibliographic reference in either service is not identical to the percentage of journals that are listed as being covered by a service. Therefore, the central question is posed: How well do the services really perform; i.e., how many "hits" are produced when the lists are

TABLE 3

COMPARATIVE PERFORMANCES
OF SSCI AND SSHI

	SSCI Hits	SSHI Hits
List I (45 items)	33 (73%)	16 (35%)
List II (38 items)	25 (65%)	10 (26%)

searched in each of the indexes? The results of this test are shown in table 3.

The accuracy of the references as listed in each service was excellent. In only one case was there a discrepancy between the original journal article and its representation in the index. (SSCI had a slightly different spelling for an author's name, probably as a result of a policy on abbreviations.) In list I, thirty-three of the forty-five items were found in SSCI, while only sixteen of the forty-five were found in SSHI. In list II, twenty-five of the thirty-eight items were picked up by SSCI; while only ten were listed in SSHI.

In light of the very different numbers of periodicals covered by the two services, recalculations were made for the number of hits that could be expected if only those journals covered by each index were used as the base number (see table 4).

SSCI slightly improves its performance on list I (78 percent) but does considerably better on list II (83 percent). The same is true for SSHI, but even so, a searcher has only an even chance of turning up a reference in the latter. Therefore, even if the searcher remembers the name of the journal in which the article originally appeared and consults the list of journals covered by SSHI before beginning the search, he or she cannot be very confident of finding the desired reference.

The time factor must be introduced to explain why a reference in a journal covered by an index does not appear there. Time lags were calculated for each reference that was found. Where the journal issue in

TABLE 4

POSSIBLE HITS BY SSCI AND SSHI

	List I	List II
SSCI	78%	83%
	(33 hits/42 possible)	(25 hits/30 possible)
SSHI	48%	55%
	(16 hits/33 possible)	(10 hits/18 possible)

which the article appeared was dated spring, fall, etc., the mid month of the quarter was assumed to be the publication date. Here the difference between the two services is readily apparent. The average time lag between publication of an article and its appearance in an index was 2.7 months for *SSCI* and 5.5 months for *SSHI*. In other words, it took a little more than twice as long for a reference to show up in *SSHI* as in *SSCI*.

One of the costs to be considered in comparing services is the psychological cost of time to searchers. How much effort is needed to garner the fruits of their labor? *SSCI* and *SSHI* are compared on this basis in table 5.

TABLE 5
MINUTES PER HIT IN *SSCI* AND *SSHI*

	List I	List II
<i>SSCI</i>	1.36 min./hit (45 min./33 hits)	2.92 min./hit (73 min./25 hits)
<i>SSHI</i>	3.25 min./hit (52 min./16 hits)	10.1 min./hit (101 min./10 hits)

There is a factor of seven between the worst and the best showings. It is relatively easy to search *SSCI*, possibly because only one or two volumes need to be consulted to get a hit, whereas the numerous issues and their cumulations make the *SSHI* search very time consuming. All searches were conducted by the same person, a professional librarian, and searching was limited to a set number of references at each session to offset the effect of fatigue.

An effort was made to translate these times into economic costs. Assuming a rate of \$9/hour, the cost of searching *SSCI* was 20.5 cents per minute for list I and almost 44 cents per minute for list II. The cost of searching *SSHI* was almost 49 cents per minute for list I and \$1.51 per minute for list II. For list I, it costs more than twice as much (238 percent) to search *SSHI* as to search *SSCI*. The gap is even greater for list II, for which the cost of searching is more than three times as much (310 percent) for *SSHI* as for *SSCI*. Combining these figures, the average cost of searching *SSHI* is 2¾ times as much (274 percent) as the cost for searching *SSCI*.

In this study, subscription rate was not

considered to be a determining factor when the effectiveness of the two services was compared. A very cheap service that fails to provide for the needs of its users is no bargain, while an expensive service that does meet these needs may be worth every penny it costs. *SSCI* and *SSHI* are here compared only on how well they perform on author searches, and no overall conclusions should be made on their effectiveness for other uses.

Summary

For author searches, *SSCI*, a computer-produced index service, outperforms *SSHI*. The time-lag factor is a particularly important element in *SSCI*'s superior showing in these tests.

The question of just how valuable timeliness is to scholars in the social sciences and humanities has not been broached. If timeliness is as important to researchers and teachers in these fields as in the natural sciences, *SSCI* should be the first place to look for bibliographic information. If timeliness is not important, it is possible that some of the missing items that should appear in *SSHI* will eventually show up there. However, *SSCI*'s greater coverage of journals gives it an advantage that is hard to match and remains the first choice for anyone doing an author search.

CONCLUSION

The two studies considered together tend to support the thesis that traditional subject-oriented index services are not particularly well suited for the retrieval of specific items. If the user looks to the library's reference tools to a large extent for this type of memory assistance, and research has shown that at least for faculty and graduate students this is true, the provision of more efficient, more mechanical services will be beneficial. Once again, it should be noted that this is not an either/or choice. No determination has been made on the relative effectiveness of the indexes for subject searching, although a large body of research literature is devoted to this topic. Each library will have to consider for itself the proper mix of services, taking into account its own assessment of the needs of its users and limitations of its budget.

REFERENCES

1. H. F. Dammers, "The Economics of Computer-Based Information Systems," *Journal of Documentation* 31:38-45 (1975).
2. Ben-Ami Lipetz and Peter Stangl, "User Clues in Initiating Searches in a Large University Catalog," *Proceedings of the American Society for Information Science* 5:137-39 (1968).
3. Allen Kent, "Generation, Use, and Transfer of Information," *Information News and Sources* 7:306-10 (1975).
4. Lipetz and Stangl, "User Clues in Initiating Searches."
5. Carl M. White, *Sources of Information in the Social Sciences*, 2d ed. (Chicago: American Library Assn., 1973).