

---

## Research and Information Retrieval Among Academic Researchers: Implications for Library Instruction

STEPHEN K. STOAN

---

### ABSTRACT

THE QUESTION OF LIBRARY instruction for faculty is linked to the issue of whether they use the library effectively in carrying on their research. Librarian-generated studies have tended to concentrate exclusively on the frequency of use of reference sources—principally abstracting and indexing systems—largely ignoring the broader intellectual and social context in which scholars function as information-generating and information-seeking individuals. Nonlibrarian researchers have been principally interested in the overall communication processes within the research community analyzed as a social system. Both bodies of literature are generally agreed that researchers do most of their information gathering using a variety of informal techniques that cause them to bypass the formal apparatus of the secondary literature or consult it only on occasion. A third body of literature that has sought to analyze the effectiveness of indexing systems has generally concluded that, for scholarly purposes, they leave much to be desired. They are incapable of incorporating the perspectival dimension needed by the scholar seeking to do original or creative work in a field. One can therefore make a strong case that the information-seeking behavior of scholars is both logical and successful given the nature of the intellectual work they are doing and the limitations of the access literature. The most extensive efforts at education have thus far failed to bring about a change in the behavior of researchers and are unlikely to do so in the future.

## INTRODUCTION

The question of library instruction for faculty is linked to the issue of whether they use the library effectively in carrying on their research. The literature on this issue has been somewhat bifurcated. On the one hand, there is a body of literature generated by librarians which "has been devoted to user studies that treat selected types of information-acquiring, or input, behavior as isolated phenomena and assume that these phenomena can be studied with little or no concern for any interactions with other communication activities" (Orr, 1971, p. 143). The surveys done by librarians have been largely unidimensional, concentrating on ascertaining the frequency with which faculty use bibliographies, indexes/abstracts, or databases that fall within the purview of reference librarians. Findings that faculty make use of a variety of other techniques to satisfy their bibliographic needs tend to be interpreted negatively and seen as justification for enhanced instructional activity designed to increase use of access tools, though what level of use would be considered appropriate or optimal is never clearly defined.

A second body of literature has been produced by nonlibrarian researchers interested in scholarly communication as a social system. This approach is more holistic, emphasizing both "informal" and "formal" channels of communication and information retrieval. It also distinguishes between "regular needs," which must be satisfied on an ongoing basis, and "episodic needs," which occur occasionally (Orr, 1977, pp. 154-55). In this scheme, structured consultation of library reference sources, including indexes, falls into the formal and episodic categories. To the extent that researchers in scholarly communication pay attention to structured literature searching, they maintain that scholars do it as needed. The success of their research speaks for itself.

The purpose of this article is, first, to offer an integrated conceptualization of the principal findings of the two bodies of literature alluded to. The second purpose is to analyze these findings in the light of yet a third body of literature, one which has sought empirically to interpret the effectiveness of information retrieval from indexing/abstracting systems and databases. From this analysis, it is hoped that one can draw some conclusions about the usefulness of library instruction for university faculty.

One precautionary note. The scope of this discussion is so limited and the body of literature that could be cited so vast that there can be no pretense of comprehensiveness. Since it is impossible in this limited space to do justice to the many nuances the literature reveals,

it has been necessary, as in all research, to exercise creative judgment in selecting the literature to be cited while synthesizing at a crude level the essential points that contribute to the argument.

### SCIENCE RESEARCHERS

By the late 1970s it was possible to report more than a thousand studies of user behavior and use of information systems carried out during the previous thirty years. The majority dealt with scientists and engineers. A dominant theme that has emerged in these studies is the occasional use of abstracting and indexing systems (Ford, 1977, p. 29). Styvendaele (1977) found that scientists and engineers at the University of Antwerp identified only 15.5 percent of their periodical citations through formal bibliographic tools, a figure consistent with a number of previous studies he cited.

More recently, Rowland (1982) reported the results of a three-year study concluded during the previous year among academic and industrial scientists in Britain. He found that the principal method of maintaining current awareness was the physical scanning of current journals. In undertaking retrospective retrieval, the main technique was footnote tracing from current literature followed by footnote tracing from a review article. Though scientists considered the abstracting journals important and were unenthusiastic about eliminating them in favor of online access only, they did not report regular consultation of them. Of those 581 scientists responding, 297 indicated that they used them once a month or less or not at all. Only seventy-eight reported weekly use. Though the scientists were favorably disposed toward online retrieval, only 223 reported ever having had a search done while 365 had not.

The rise of online access has caused many to assume that it would ultimately revolutionize the way scientists and other researchers go about doing their work. An early study (Knightly, 1979) designed to see if computer searching was replacing the use of other techniques of information gathering concluded that it was not. Online was used as an occasional supplementary approach to more traditional techniques. A more elaborate study by Bayer and Jahoda (1981) of 262 industrial scientists and 70 university chemists loaded the dice in favor of online retrieval by providing unlimited free access to a search service for a year. The authors learned that use of the service did not diminish the total amount of time devoted to information-gathering activities and did not cause scientists to discontinue more traditional approaches. They found no significant impact on "continuing use of traditional information retrieval strategies" (p. 328). Pre- and post-tests revealed that the more frequent users of the

service even ended up significantly increasing "their assessment of the utility of scanning primary sources and use of citations from other works" (p. 329).

Borgman, Case, and Ingebretsen (1985) conducted a survey with only a 19 percent response rate on faculty use of database searching in six departments at ten universities. Of this sample, 41 percent reported never having a database search done. The commonest response to a frequency question on online searching was once a year, with only a quarter of the sample reporting use more frequent than three times a year (pp. 311-13).

More recently, Horner and Thirlwall (1988) reported on a survey of faculty at the University of Manitoba designed to ascertain their use of both mediated and personal searching. Among scientists and engineers, only 7 percent reported having a search done for them many times. Another 31.9 percent reported occasional use, 24.3 percent reported rare use, and 36.8 percent indicated no use at all. The comparable figures for personal searching were, respectively, 5.8, 12, 13.5, and 68.7 percent. This survey is especially valuable because it demonstrates that exploitation of end-user systems is not a reason for the occasional use of mediated systems.

Bichteler (1986) made an effort to identify geoscientists who were doing their own searching since a number of studies reveal relatively occasional use of mediated services by geoscientists and even a somewhat negative attitude about the overall value of computer retrieved bibliographies. Her national search tracked down only a small number of end-users, many of whom were practitioners rather than academicians. They typically did five to ten searches a year, many of which were author searches, rarely consulted a thesaurus or other documentation, and seldom did large retrospective searches (pp. 46-48). Torok and Hurych (1986) surveyed science, social science, and humanities faculty in twenty universities in an effort to determine level of interest in end-user searching. Though expressed interest was high, only 33 percent reported wanting as many as three searches a year (p. 337). These figures bespeak both a low level of demand for online retrieval and a limited market for end-users, especially adept ones.

A judicious assessment of the impact of online retrieval on the conduct of scientific research is given by Orr, who noted that scientists seldom request more than one or two searches a year even where they have unlimited free access to search services, "as numerous systems designers have learned to their dismay when the computer search systems they installed were utilized to only a fraction of design capacity" (p. 160). Once the novelty has worn off, Orr concluded, scientists tend to use the system only for essential needs.

The literature generated by students of scholarly communication in the sciences has, as might be surmised, come to similar conclusions about the sporadic use of formal bibliographic retrieval systems, whether print or online, by scientific researchers. Two excellent and well documented summaries are the works edited by Nelson and Pollock (1970) and Garvey (1979). The main thrust of these studies has been to identify both informal and formal information retrieval channels among scientists, with the former predominating. The much mentioned "invisible college" (Paisley, 1965; Crane, 1972; Cronin, 1982; Chubin, 1983) is a reality, though there appear to be many of them at any one time, depending on the discipline, subdiscipline, and nature of the area being researched. Some research fronts have strong invisible colleges; others have nearly nonexistent ones. They are somewhat evanescent in character. Still they provide a mechanism through which scientists in a specialized area of research may carry on routine communication through personal contacts at conferences and symposia, and the exchange of conference papers, technical reports, preprints, and reprints that precede the appearance of the reported research results in a refereed journal. Though librarians have tended to think of the journal article as "new" and its appearance in abstracting journals or reviews of research as "old," for large numbers of scientific researchers the journal article itself is old.

The more important scientists in a field have particularly well developed informal communication networks. They are the people who obtain large grants, invite others to share research projects with them, serve as officers on important scholarly organizations and as editors of leading journals, referee grant proposals and journal manuscripts, organize symposia, select students for admission to graduate school, write recommendations for other people's research proposals, and so on (Garvey, 1979, p. 12). In this way, they are strategically situated to keep current on research being carried on by many researchers in their fields. It is worth noting, too, that a small number of individuals account for a high percentage of published research in the scientific fields. Griffith and Miller (1970) state that younger researchers, who are not yet established, seem to make more use of structured communication channels primarily because they are still outside the informal networks (pp. 134-35).

These informal channels, as Garvey (1970) points out, satisfy a number of psychological and practical needs for scientific researchers. They offer an opportunity prior to publication to work flexibly on the project, to present the results tentatively so as to reshape it based on feedback from others, to obtain reinforcement from (and commiserate with) kindred spirits, particularly if one is part of a network seeking to challenge a current orthodoxy, to establish

hitherto nonexistent but potentially valuable contacts with important researchers, and to control those with whom one exchanges information (pp. 143-45, 153-56).

### SOCIAL SCIENCE RESEARCHERS

The discipline of psychology serves as the bridge between the sciences and the social sciences, evincing characteristics of both areas. One of the major studies of scientific communication, carried out by the American Psychological Association (APA, 1963, 1965, 1969), is frequently cited in both science and social science communications research. Though there had been isolated studies by librarians and others of the information-seeking behavior of social scientists in other disciplines, the first and only large-scale study was that undertaken at Bath University in the late 1960s under the direction of Maurice Line (*Information Requirements*, 1971). Useful syntheses of the principal findings can be found in Line (1971), Evans (1974), Morrison (1979), Stoan (1986), and Slater (1988).

In broad outline, social scientists rely heavily on citations identified in book and journal literature, on recommendations from colleagues, on personal collections and bibliographic files, and, in the more book-oriented disciplines, on browsing. Their overall use of indexing systems may be lower than that of scientists. These findings have been corroborated by smaller studies conducted by Wood and Bower (1969), Styvendaele (1977), Stenstrom and McBride (1979, 1982), and Stieg (1981). More recently, Folster (1989) studied faculty in four social science departments finding once again that reading journals in their own field, tracing references, consulting personal collections, consulting colleagues, scanning journals in other fields, and attending conferences all ranked above use of abstracts or indexes as techniques of information gathering. Use of online searching ranked at the bottom. Thaxton (1985) found in a study of faculty and graduate students in psychology at Georgia State that informal communication patterns were indeed strong and that formal use of access tools was modest. Students who had attended instructional sessions offered by the librarians did not differ in their behavior from those who had not.

Horner and Thirlwall's (1988) study at the University of Manitoba revealed that only 8.8 percent of social science faculty reported using mediated online searching many times, 29.9 percent occasionally, 22.4 percent rarely, and 39 percent never. Comparable figures for personal searching were, respectively, 6.9, 15.1, 14.2, and 64.8 percent. These rates of use were almost identical to those of scientists and engineers. Garvey, Lin, and Nelson (1970) generalize that the informal

communication channels described earlier for the sciences exist also in the social sciences, but that they are generally less tightly structured, more unpredictable, and work more slowly (p. 297).

The sporadic use of indexes by social scientists has been traced to a number of factors. Indexing systems in the social sciences are numerous, small, cover limited numbers of journals, often selectively, fail to include book literature, which is heavily used by most social scientists, and encounter major problems in terminological control, which is in no way resolved by keyword access (Evans, 1974; Stoa, 1986, pp. 10-11). Moreover, indexing tools in the social sciences cannot compensate for the very powerful eclecticism and multidisciplinary use of materials evinced by most of the social sciences (Earle & Vickery, 1969; Broadus, 1971; Line & Roberts, 1976; Brittain, 1979; Line, 1971, 1980, 1981). Lastly, social scientists often rely for primary data on a wide variety of materials such as collections of printed documents, archival materials, statistics, government publications, polling results, memoirs, speeches, autobiographies, newspapers, eyewitness accounts, etc., that are not indexed as social science literature *per se* and require a great deal of creative ingenuity on the part of the researcher to conceptualize as significant and then track down.

## HUMANITIES RESEARCHERS

For the most part, there is less empirical evidence about how humanists work than about other scholars. The most systematic studies have emerged from the Centre for Research on User Studies in Sheffield, England, which set out to do for humanist research what Bath did for social science research. An early study reported by Corkill (1978) involved a mail survey to professors and graduate students in English, French, history, music, and philosophy at thirty-five universities in the United Kingdom. The 612 responses from faculty represented a 64.4 percent response rate. Corkill reported that humanists used personal collections heavily, generally consulted a very broad range of materials, many of which were quite old, made heavy use of book literature, consulted a great deal with other scholars, did little delegation of information retrieval, did the basic research work alone, and relied on libraries heavily for much of the material used (pp. 55-58). In terms of information-seeking techniques, more than 90 percent reported using publishers' catalogs and inspection copies of books, scanning current journals, following the library's accession lists, and maintaining informal contacts with colleagues and other researchers. More than 80 percent obtained information at conferences and through subscriptions to key journals in their fields. Only 21.6 percent used bibliographies or abstracting/indexing systems (p. 84).

Stone (1982) added the following points about humanist scholars. They seldom collaborate but do consult with others a great deal, browse very heavily in the stacks since they are often not seeking anything specific, utilize a variety of research methods and materials, only occasionally consult secondary services, which, as in the social sciences, index only journal literature, emphasize personal opinion and interpretation very heavily, and work best in an open stack arrangement with a large monograph collection arranged by subject classification.

Weintraub (1980), Garfield (1980), Broadbent (1986), Fabian (1986), and Wiberley and Jones (1989) have corroborated this general picture, emphasizing the lack of an identifiable "research front" or cumulativeness in humanities research. Humanities literature could perhaps be better described as aggregative rather than cumulative, for one can still reinterpret Plato the same as Alfred North Whitehead, or John Donne the same as James Baldwin. Older literature is continually consulted anew, and there is little compulsiveness about "current awareness" since any one interpretation can be considered as valid a contribution as another.

With regard to use of online retrieval systems, Horner and Thirlwall's (1988) indicative study at Manitoba found that 1.9 percent of humanities faculty used mediated search services many times. Only 9.4 percent used such services occasionally, 13.2 percent rarely, and 75.5 percent never. Comparable figures for personal searching were 5.7 percent, 9.5 percent, 11.4 percent, and 73.3 percent. These figures are lower than for the sciences and social sciences and indicate that databases that search only journal literature, and go back only a decade or so, are of limited use to scholars who rely heavily on books that may go back many decades.

A very critical aspect of humanities research, also valid for much social science research, is the unique nature of the monograph and even the journal article. Whereas in the sciences a journal article reports the results of one's research, in the humanities and social sciences a monograph or journal article is the result of one's research.<sup>1</sup>

Wilson (1980) summarizes this with his accustomed insights, using historians as an example:

historians accumulate bodies of fact and also accumulate competing explanations and interpretations of the facts without apparent limit. The historian's results are not conclusions which can be stated briefly and impersonally and recorded in a reference book of historical findings. The monograph the historian writes does not simply present his results; it is itself the result. It is a piece of art, of high or low quality, and the art cannot be factored out. (p. 12)

Farther on, Wilson notes that syntheses or abstracts of social science (and humanities) literature are unsatisfactory substitutes for the originals. "To know what social scientists have done, one has to read their works, for their works are what they have done" (p. 18).

This view of the monograph as a work of personal creativity, not unlike a musical composition or painting, causes humanists to be able to identify hundreds of scholars and the monographs they have written in much the way that a student of music can identify hundreds of musical compositions and a student of art hundreds of individual works of art. Though humanists often write articles—which may themselves become minor classics as works of unusual insight and influence—their articles are often only a way station on the road to a fuller exposition in monographic form; hence the greater emphasis on the monograph as offering the most comprehensive treatment of the scholar's fully developed ideas. This same statement can be made of social researchers, like political scientists, who work in book-oriented disciplines.

#### PERSONAL COLLECTIONS AND LIBRARY COLLECTIONS

Before leaving the discussion of what we presently know of how scientists, social scientists, and humanists work, it would be useful to include a few paragraphs on the development and use of personal collections. Soper's (1976) extensive survey of 178 faculty in the sciences, social sciences, and humanities revealed that 98 percent of them had personal collections of considerable size that existed regardless of the size of library they had access to. These collections included books, journal subscriptions, copies of papers delivered at conferences, photocopied articles, preprints, reprints, government publications, research reports, and so on. Essentially, researchers seek to build up a library of materials focused on their principal research interests and use the institutional library as a supplement for more expensive, rarer, less frequently needed, or more diffuse subject-oriented materials.

Soper (1976) determined that faculty tend to consult their own collections first, then those of their colleagues, then a departmental collection (if one exists), and lastly the institutional library. All surveyed faculty ranked the importance of their personal collections as first (humanities and social sciences) or first or second (sciences) in their research (pp. 397-401). Their primary reason for developing personal collections is the convenience of the immediate accessibility of much desired or frequently consulted items, since faculty see access to the library's materials as uncertain and inconvenient.

Soper (1976) also determined that, in their own research, faculty cite materials from their personal collections a great deal. The

scientists and social scientists she studied cited personally owned materials about 73-74 percent of the time. Humanists, more dependent on a huge and very diffuse monographic literature, cite their personal collections only 36 percent of the time (pp. 402-13). Since scholars set out to accumulate materials that bear most directly on their areas of specialized research interests, it is not surprising that their personal collections can assume such significance. It is also not surprising that these collections can provide access to a wealth of focused bibliographic information.

To point out the significance of personal collections in scholarly research in no way diminishes the importance of library resources, which all faculty report to be valuable for their research and teaching. The research library fills an indispensable role by acquiring and maintaining rarer, more expensive, and less commonly needed materials that serve as supplements in research and assist in maintaining general currency in the discipline for teaching purposes. "Insofar as the . . . library houses copies of information sources that figure in one's reserve supplies of information, to be consulted in case of need, the library provides a benefit that is independent of the actual frequency with which the sources are consulted" (Wilson, 1977, p. 85).

#### AN INTERPRETATION OF RESEARCH FINDINGS

In 1981, Wilson observed in an article on user studies that research in this area has suffered from concentrating on the "*means* by which people discover information (often analysed in terms of the information researcher's view of how the user *ought* to have been seeking information) rather than upon the *ends* served by the information-seeking behaviour" (p. 10). The result of this bias, Wilson observes, is dissatisfaction with the results of the studies, "since the service implications have been far from clear" (p. 10). One must seek to understand the psycho-social context of information seeking in order to understand what information a person wants, why he or she wants it, and what techniques she or he chooses for obtaining it. Though Wilson was describing all user study research, his trenchant observations certainly apply in the case of studies of scholars.

The picture that emerges from an examination of the literature on faculty information gathering can be conceptualized in several different ways. At the most basic level, it can simply be stated that informal techniques for keeping up with the literature and retrieving materials useful for research prove to be satisfactory to scholars. Hence their lack of concern about changing personal behaviors that for

them have been successful. To put it another way, researchers do not see a problem in terms of bibliographic retrieval and so are not seeking a solution.

At another level, one can note, as Orr (1970) has, that a good way to interpret researcher behavior is through holistic analysis of all options open to the researcher for retrieving bibliographies or other data. If the individual must allocate limited amounts of time and energy toward information seeking, and that information can be obtained satisfactorily in a number of different ways because of the considerable redundancy built into the system through the invisible college, direct consultation, personal collections and bibliographic files, references and footnotes, or bibliographies, indexes, and abstracts, researchers opt for those techniques that have the highest reward-cost ratio while offering the greatest psychological gratification by serving a number of needs simultaneously. In this situation, "their observed preference for informal channels is completely understandable" (p. 155). Orr also cites evidence that the strongest single predictor of publication productivity is the amount of informal contacts with other researchers (p. 168).

Yet another way of conceptualizing the approach used by scholars is to note that it emphasizes information-retrieval channels that offer guidance from other experts in their fields, whether in the form of informal communication through invisible colleges, consultation with colleagues, scanning newly published literature for current awareness purposes, or paying close attention to the literature cited by other scholars in their monographs or articles. In so doing, researchers are obtaining scholarly analysis and guidance from their peers, who provide the intellectual context indispensable for understanding research in the discipline. The numerous surveys showing that consultation with librarians ranks very low as a means of information retrieval for faculty are further evidence that librarians, not being viewed as experts in subject disciplines, are outside of the research loop in any fundamental sense.

The emphasis on information-retrieval techniques that link researchers directly to the ideas, interpretations, suggestions, comments, and views of their peers dovetails neatly with the sizable literature on the intellectual processes involved in research. These studies point to the powerful influence of creative insight and intuition that come only from a well instructed mind working continually with the subject matter of the discipline.<sup>2</sup> They emphasize that, despite the popular conception of the "scientific method" learned in grade school, research is normally random, nonlinear, and nonsequential.

Since consulting the literature is but one dimension of a complex intellectual process of ongoing dialog with the subject content of the discipline as the research project germinates, evolves, matures, and bears fruit in the mind of a researcher, it is difficult to generalize where a structured literature search “fits,” if at all, in the execution of any particular project in any particular discipline. Just as the intellectual processes involved in research are often random, exploitation of library materials is also random (Grover & Hale, 1988, p. 11). The evolution of the project in the mind of the researcher “dictates the sources sought out at each stage along the way. A new idea generated from one source, an original insight springing from another, may alter the direction of the quest and the kind of material being sought” (Stoan, 1984, p. 102).

The preceding observations suggest that there is a defensible logic in the information retrieval techniques used by scholarly researchers, who would accept them as self-evident on an experiential level. Still, these observations do not in and of themselves “prove” the superiority of informal approaches on an empirical level. Is it not possible, as many librarians suggest, that more routine consultation of the secondary services will improve the quality of research? In one sense, the answer to this question is immaterial. Since the faculty are part of an elaborate social system with its own rewards and punishments, they only respond to penalties imposed by their own peers—other researchers in their fields. Since librarians are not part of their social system, faculty are largely unconcerned with their perceptions.

In the interests of scholarly objectivity, however, it may be possible to quantify, in a backhanded way, at least one dimension of this problem. There have been a number of studies aimed at evaluating the effectiveness of information retrieval from bibliographic services, both print and online. These would give at least some indication of the likely usefulness of organized bibliographies to a researcher.

Three splendid articles by Swanson (1977, 1986a, 1986b,) should be considered essential reading by all librarians seeking to understand the intellectual difficulties inherent in all information retrieval. Though it is possible to present in synopsis form some of Swanson’s major data and arguments, there is no adequate substitute for reading the originals.

An underlying theme in Swanson’s articles is that there is a vast body of public knowledge contained in the written record whose retrieval is highly problematical and always incomplete because there is no satisfactory way of labeling (indexing) each document for purposes of retrieval for every possible use to which it might be put. Indexing terms merely indicate the main thrust of the article or book

as viewed from the perspective of the author. They can never account for other perspectives from which that document, or even a part of that document, might be used by other researchers working in essentially unrelated areas. Information contained in books or articles may be used in the future in ways as yet undreamed of in the present.

Indexing systems, essentially, cannot overcome the perspectival problem inherent in all retrieval. Since the only way to guarantee that no potentially useful, or relevant, information might be missed would be to examine the entire written record of the human race—an obvious impossibility—one must conclude that all information retrieval is essentially incomplete, “or, if it were complete, we could never know it. Information retrieval, therefore, is necessarily uncertain and forever open-ended” (Swanson, 1986a, p. 114). In doing bibliographic retrieval, we are looking for what we do not know and are never certain how much we have not found.

To illustrate with a purely hypothetical example, let us suppose that a historian is doing research on the Indian removal policy of Andrew Jackson. A standard literature search using the obvious indexing terms will only turn up documents whose main thrust is Jackson’s Indian removal policy during the 1830s. It will not turn up an article on the agrarian economy of Alabama from 1820 to 1850, which contains a splendid, well documented two pages on the economic impact of Indian removal on this part of the South. It will not turn up an article on changing interpretations of the U.S. Constitution from 1801 to 1861 which may contain a fine page on the constitutional implications of Indian removal. It will not turn up the diary of a white settler who obtained land during this period. It will not turn up an autobiography of one who spent two years of his life surveying lands obtained from the Indians. It will not turn up a book on British foreign policy in the nineteenth century that may contain information on international reactions to Jackson’s policy. It will not turn up a monograph on the history of antebellum Mississippi containing a chapter on Indian removal. It will not turn up a general history of the Cherokee people.

Though all of these other pieces of published literature would be useful, each has been indexed in ways that no bibliographic searcher could possibly have thought to include in a search strategy. Only the creative mind of a highly knowledgeable researcher drawing bits and pieces of data from a wide range of sources and obtaining leads through numerous channels would likely come across all of these items and weave them together into a unique work of scholarship. In the last analysis, it is the mind of the researcher that endows a document with “relevance” by conceiving a way in which it, or even a small part of it, fits into his/her emerging research scheme.

A descriptor assigned by another party promising that the "topic" of a document conforms to the general area of one's own research does not guarantee relevance. Indeed, this is the heart of the information retrieval dilemma. Documents indexed under the obvious subject terms may prove to be irrelevant; documents indexed under subject terms one would never think to look under may prove to be relevant. As if this were not enough, the researcher must somewhere make an informed judgment that she must cease gathering data and commit ideas to paper, even though there may still be much unexploited data of whose existence she is unaware.

In bolstering his arguments on the inherent limitations of information retrieval systems, Swanson surveys four experiments effected in both manual and automated environments that demonstrate with ineluctable empirical evidence the inherent limitations of the topical "literature search." In 1953, Documentation, Inc. and the Armed Services Technical Information Agency Reference Center challenged each other to a contest to test the relative effectiveness of their indexing systems. After teams from the two agencies had searched for bibliographic information on ninety-eight search requests from a pre-selected set of 15,000 documents held in common, they learned that one team had retrieved 2,220 documents and the other only 1,560. Most disturbing is that they had retrieved only 580 items in common. After reviewing the complete set jointly retrieved, the two teams could only agree on 1,390 items as being relevant to the ninety-eight search requests. They disagreed on 1,577 items (Swanson, 1986b, pp. 389-90).

The Aslib Cranfield Project in England carried out a series of information retrieval experiments in 1966 which, Swanson calculated, missed about 92 percent of the potentially relevant documents in the bibliographic base used. The MEDLARS test of 1967, which sought to test both precision and recall for the new automated system by comparing retrieval on 302 questions to a pre-identified list of relevant articles resulted in an average precision rate of 50 percent and an average recall rate of 58 percent. These figures, being averages, concealed the fact that precision and recall on any individual question searched fell almost uniformly anywhere on a scatter chart, making it impossible to make any predictive claims for the precision or recall rates of a computer search. Lastly, the SMART-MEDLARS comparison experiments, carried out in the late 1960s and early 1970s, were unable to demonstrate empirically that free-text searching was more successful than the already problematical controlled vocabulary searching in computerized literature retrieval (Swanson, 1977, pp. 131-36).

Vincent (1984), writing specifically of research in the humanities, also noted the limitations of access tools in helping a scholar capture a new insight or interpretation: "not only do indexes and computer databases continue to have limitations as pathfinders to scholarship," he wrote, "they are utterly incapable of placing an idea or concept into its proper context . . ." (p. 181).

More recently, Weinberg (1987) examined the issue of indexing in the online environment, coming to much the same conclusions as Swanson. She stressed that indexing terms relate only to the "aboutness" of a document, whereas scholars, seeking to solve problems or observe data in new ways, are interested in "aspect," an area in which indexing systems, even with free-text capabilities and abstracts, fail totally.

Truly striking evidence of the essential incompleteness and utter unpredictability of subject retrieval online has been supplied by Trivison, Chamis, Saracevic, and Kantor (1986). They obtained forty search questions, assembled a group of thirty-six experienced online searchers, and submitted each question to groups of five searchers based on areas of subject expertise. Each searcher then independently conducted an online search in a designated DIALOG database to guarantee complete comparability of retrieval results. The four organizers of the project also contributed searches. The nine bibliographies retrieved for each search question were merged into a "union of output" to be submitted to the requestor for evaluation as to relevance. This union of output was used as the base upon which to calculate the precision and recall ratios for each search run.

The results were striking. In preliminary results reported on five questions, the precision for the unions of output varied from less than 20 percent for one question to more than 90 percent for another. The numbers of relevant documents retrieved by individual searchers on the same question varied from one to twenty-seven on one question, nine to forty-one on another, one to forty-three on a third, zero to fifty on a fourth, and four to forty-four on a fifth. Recall by individual searchers was uniformly low. Only one searcher found more than half of the documents collectively identified by the group. Others ranged from 0.0 percent to 48.1 percent, with most falling under 30 percent. There was little overlap among searches on the same topic, and the search strategies and numbers of commands and search terms used varied widely.

The only conclusion one can come to based on these results is that one can legitimately make very few claims for online searching. Depending on the search problem and the searcher, results can vary widely. It would be rash to talk about an "exhaustive" or

“comprehensive” search. A searcher could make no promises with regard to an average level of either recall or precision. At best, one can only promise the patron to find something that might be a useful lead.

In a similar vein, the UK Scientific Documentation Centre, with funding support from the British Library, conducted a three year study in the mid 1980s of bibliographic retrieval in science and technology with a view to identifying the most effective techniques of information retrieval. The extensive study concluded that online searching was the least effective method of bibliographic retrieval, being systematically outperformed by print sources by a wide margin (Davison, et al., 1988).

There is certainly much evidence that bibliographic retrieval from secondary sources, whether print or online, is imprecise, tentative, inconclusive, and incomplete. The empirical studies are the more persuasive since nearly all of them have dealt not with social science and humanities literature, whose “soft” terminology makes retrieval more problematical, but with the sciences, whose terminology has supposedly been “harder” and more precise. The studies confirm what researchers have long observed at an experiential level—namely, that the best of bibliographies, indexes, and database searches are merely sometimes helpful supplements to other methods of bibliographic retrieval. As a practicing social scientist (Rush, 1974) in the United Kingdom put it, “not only is the computer severely handicapped, even helpless, when faced with terminological inexactitude, but so also is the bibliographer, the indexer and the abstractor” (p. 94).

#### INSTRUCTION FOR THE FACULTY

The conclusions to be drawn from the evidence presented are these. The faculty rely on a wide variety of information-retrieval techniques, many of them informal and most of them geared toward obtaining some kind of expert guidance from other scholars as part of the retrieval process. The nature of the research process is such that their need to carry out structured literature searches, whether in print or online, is occasional. Their experiences with bibliographic tools, including online ones, have evidently not been so positive as to convince them that more frequent exploitation of these sources would significantly benefit their research. Indeed, structured literature searches in print or online sources, using either assigned descriptors or free text capabilities, can be shown empirically to suffer from grave limitations in terms of precision and recall. Overall, faculty are generally satisfied with the way they are carrying on their research and doing literature retrieval for research purposes. Their behaviors

in this regard have been "successful." One result is that at least a generation of efforts on the part of librarians to reorient faculty behavior through education programs of some kind or another have changed nothing.

The Bath University researchers initially reacted to evidence of low or occasional use of bibliographic sources as indicating a need for expanded instruction. Evans (1974), a participant in the Bath investigations, later reported that the experiment of providing social researchers with an "information officer" to retrieve information for them was successful (pp. 85-86). But efforts to offer seminars on information seeking, library style, turned up only three volunteer participants (p. 90). Brittain (1985), also a Bath participant, reported, in surveying Bath findings, that expanded user education efforts failed to change researcher behavior. In grasping for some explanation for "user resistance" to more systematic exploitation of the access tools, Brittain could only surmise that lack of cumulativeness in social science research removed the penalties for poor work (pp. 266-70).

Such an explanation, of course, could be equally well applied to research in library science itself, where the lack of paradigms, theories, theses, conceptual frameworks, and cumulativeness has long been noted. As a matter of fact, if librarians are correct that frequent and systematic exploitation of secondary services is essential to good research, it would follow that research carried out by librarians should be superior to that of other scholars, assuming, of course, that librarians practice what they preach. It would be difficult to demonstrate, however, that research carried out by librarians is consistently better than research carried out in other fields. It would also be difficult to demonstrate that library researchers do better literature retrieval, however one might measure that, than researchers in other fields.

Where does this leave library instruction for faculty? The library should certainly undertake to offer training in such areas as the mechanics of retrieving from the online catalog or from CD-ROMs. It should seek to keep faculty informed of important reference sources and new reference acquisitions that may have particular usefulness to a department or to specific faculty members. It should also try to provide new acquisitions lists of materials acquired in the general collection. But the ultimate conclusion offered by the Bath University researchers in 1971 on user education for the faculty continues to be valid:

The information profession sometimes assumes that researchers want to, and can, work in a systematic way in dealing with bibliographical material and that the bibliographical system is about the only system, or at least the most important system, for the transfer of information. In view of the overwhelming evidence that social scientists do not perform

in this way, such assumptions (sometimes followed by exhortations) should be avoided. User education may go a long way to alerting researchers to potentially useful bibliographic tools and ways of using them; but it is doubtful if it could do more. (Information Requirements, 1971, p. 91)

## NOTES

1. This unique dimension of the social science or humanities article compared to the science article helps explain a number of observed differences in the literature. Science articles, being essentially barebones descriptions of the results of a research project, are generally briefer and are produced in much greater profusion. Consequently, there are many more articles per issue, more issues per year, more journals in which to publish them, and the rejection rate is quite low. In the social sciences and humanities, by contrast, the reverse is true in each case.
2. As indicated, there is a wealth of literature generated in recent decades seeking to analyze the intellectual processes involved in scholarly research. The following list is only suggestive, but, it is hoped, helpful. See Mills, C. W. (1959). *The sociological imagination*. New York: Oxford University Press; Popper, K. R. (1959). *The logic of scientific discovery*. New York: Harper; Popper, K. R. (1963). *Conjecture and refutations*. New York: Harper; Popper, K. R. (1972). *Objective knowledge: An evolutionary approach*. London: Oxford University Press; Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science*. San Francisco, CA: Chandler; Polanyi, M. (1958). *Personal knowledge: Towards a post-critical philosophy*. Chicago, IL: University of Chicago Press; Polanyi, M. (1959). *The study of man*. Chicago, IL: University of Chicago Press; Polanyi, M. (1967). *The tacit dimension*. London: Routledge & Kegan Paul; Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press; Hammond, P. E. (Ed.). (1964). *Sociologists at work: Essays on the craft of social research*. New York: Basic Books; Watson, J. D. (1968). *The double helix*. New York: New American Libraries; Ravetz, J. R. (1971). *Scientific knowledge and its social problems*. New York: Oxford University Press; and Latour, B., & Woolgar, S. (1979). *Laboratory life: The social construction of scientific facts*. Beverly Hills, CA: Sage.

## REFERENCES

- American Psychological Association. *Reports of the American Psychological Association Project on Scientific Information Exchange in Psychology* (vol. 1, 1963; vol. 2, 1965; vol. 3, 1968). Washington, DC: APA.
- Bayer, A. E., & Jahoda, G. (1981). Effects of online bibliographic searching on scientists' information style. *Online Review*, 5(4), 323-333.
- Bichteler, J. (1987). End user database searching in geoscience. In R. A. Bier, Jr. (Ed.), *The user and geoscience information* (Proceedings of the Twenty-First Meeting of the Geoscience Information Society, 10-13 November 1986, San Antonio, Texas) (pp. 45-53). Alexandria, VA: Geoscience Information Society.
- Borgman, C. L.; Case, D. O.; & Ingebretsen, D. (1985). University faculty use of computerized databases: An assessment of needs and resources. *Online Review*, 9(4), 307-332.
- Brittain, J. M. (1979). Information services and the structure of knowledge in the social sciences. *International Social Science Journal*, 31(4), 711-728.
- Brittain, J. M. (1985). Relevance of social science output worldwide. *International Social Science Journal*, 37(2), 259-275.
- Broadbent, E. (1986). A study of humanities faculty library information seeking behavior. *Cataloging & Classification Quarterly*, 6(3), 23-37.
- Broadus, R. N. (1971). The literature of the social sciences: A survey of citation studies. *International Social Science Journal*, 23(2), 236-243.
- Chubin, D. E. (1983). *Sociology of sciences: An annotated bibliography on invisible colleges, 1972-1981*. New York: Garland.

- Corkill, C., & Mann, M. (1978). *Information needs in the humanities: Two postal surveys* (CRUS occasional paper 2). Sheffield, England: Centre for Research on User Studies.
- Crane, D. (1972). *Invisible colleges: Diffusion of knowledge in scientific communities*. Chicago, IL: University of Chicago Press.
- Cronin, B. (1982). Invisible colleges and information transfer: A review and commentary with particular reference to the social sciences. *Journal of Documentation*, 38(3), 212-236.
- Davison, P. S.; McDonagh, M.; Meldrum, J. D. P.; & Moss, A. (1988). *International bibliographic review on costs and modelling in information retrieval: Report on completion of data collection, assessment of the efficiency of searching, and on means of indexing this information*. London: British Library.
- Earle, P., & Vickery, B. (1969). Social science literature use in the UK as indicated by citations. *Journal of Documentation*, 25(2), 123-141.
- Evans, S. (1974). Information services for social scientists in the universities. In F. H. Ayres & J. Hall (Eds.), *Information services in university libraries* (Standing Conference of National and University Libraries [SCONUL], Exchange of Experience Seminar on Information Services in University Libraries) (pp. 84-92). Bradford, England: SCONUL.
- Fabian, B. (1986). Libraries and humanistic scholarship. *Journal of Librarianship*, 18(2), 79-92.
- Folster, M. B. (1989). A study of the use of information sources by social science researchers. *The Journal of Academic Librarianship*, 15(1), 7-11.
- Ford, G. (Ed.). (1977). *User studies: An introductory guide and select bibliography*. Sheffield, England: Centre for Research on User Studies.
- Garfield, E. (1980). Is information retrieval in the arts and humanities inherently different from that in science? The effect that ISI's citation index for the arts and humanities is expected to have on future scholarship. *Library Quarterly*, 50(1), 40-57.
- Garvey, W. D. (1979). *Communication: The essence of science*. New York: Pergamon.
- Garvey, W. D.; Lin, N.; & Nelson, C. E. (1979). Communication in the physical and social sciences. In W. D. Garvey (Ed.), *Communication: The essence of science* (pp. 280-99). New York: Pergamon.
- Griffith, B. C., & Miller, A. J. (1970). Networks of informal communication among scientifically productive scientists. In C. E. Nelson & D. K. Pollock (Eds.), *Communication among scientists and engineers* (pp. 125-40). Lexington, MA: Heath Lexington Books.
- Grover, R., & Hale, M. L. (1988). The role of the librarian in faculty research. *College & Research Libraries*, 49(1), 9-15.
- Horner, J., & Thirlwall, D. (1988). Online searching and the university researcher. *The Journal of Academic Librarianship*, 14(4), 225-230.
- Knightly, J. J. (1979). Traditional information gathering vs. the computer: Surprising findings. *Tennessee Librarian*, 31(3), 32-38.
- Line, M. B. (1971). The information uses and needs of social scientists: An overview of INFROSS. *Aslib Proceedings*, 23(8), 412-434.
- Line, M. B. (1980). Secondary services in the social sciences: The need for improvement and the role of librarians. *Behavioral and Social Sciences Librarian*, 1(4), 263-273.
- Line, M. B. (1981). The structure of social science literature as shown by a large-scale citation analysis. *Social Science Information Studies*, 1(2), 67-87.
- Line, M. B.; Brittain, J. M.; & Cranmer, F. A. (1971). *Information requirements of researchers in the social sciences*. (1971). 2 vols. Bath, England: Bath University of Technology.
- Line, M. B., & Roberts, S. (1976). The size, growth and composition of social science literature. *International Social Science Journal*, 28(1), 122-159.
- McBride, R. B., & Stenstrom, P. (1980/81). Psychology journal usage. *Behavioral and Social Sciences Librarian*, 2(8), 1-12.

- Morrison, P. D. (1979). Since Bath: A review of published information transfer studies in the social and behavioral sciences, 1974 through 1978. *Behavioral & Social Sciences Librarian*, 1(1), 5-22.
- Nelson, C. E., & Pollock, D. K. (Eds.). (1970). *Communication among scientists and engineers*. Lexington, MA: Heath Lexington Books.
- Orr, R. H. (1970). The scientist as an information processor: A conceptual model illustrated with data on variables related to library utilization. In C. E. Nelson & D. K. Pollock (Eds.), *Communication among scientists and engineers* (pp. 143-89). Lexington, MA: Heath Lexington Books.
- Paisley, W. J. (1965). *The flow of (behavioral) science information: A review of the research literature*. Stanford, CA: Institute for Communication Research.
- Rowland, J. F. B. (1982). The scientist's view of his information system. *Journal of Documentation*, 38(1), 38-42.
- Rush, M. (1974). The information requirements of a university social scientist. In F. H. Ayres & J. Hall (Eds.), *Information services in university libraries* (pp. 93-103) (Standing Conference of National and University Libraries [SCONUL] Exchange of Experience Seminar on Information Services in University Libraries). Bradford, England: SCONUL.
- Slater, M. (1988). Social scientists' information needs in the 1980s. *Journal of Documentation*, 44(3), 226-237.
- Soper, M. E. (1976). Characteristics and use of personal collections. *Library Quarterly*, 46(4), 397-415.
- Stenstrom, P., & McBride, R. B. (1979). Serial use by social science faculty: A survey. *College & Research Libraries*, 40(5), 426-431.
- Stieg, M. F. (1981). The information needs of historians. *College & Research Libraries*, 42(6), 549-560.
- Stoan, S. K. (1984). Research and library skills: An analysis and interpretation. *College & Research Libraries*, 42(2), 99-109.
- Stoan, S. K. (1986). Social science literature. In W. H. Webb (Ed.), *Sources of information in the social sciences* (3d ed.) (pp. 1-62). Chicago, IL: American Library Association.
- Stone, S. (1982). Humanities scholars: Information needs and uses. *Journal of Documentation*, 38(4), 292-313.
- Swanson, D. R. (1977). Information retrieval as a trial-and-error process. *Library Quarterly*, 47(2), 128-148.
- Swanson, D. R. (1986a). Undiscovered public knowledge. *Library Quarterly*, 56(2), 103-118.
- Swanson, D. R. (1986b). Subjective versus objective relevance in bibliographic retrieval systems. *Library Quarterly*, 56(4), 389-398.
- Thaxton, L. (1985). Dissemination and use of information by psychology faculty and graduate students: Implications for bibliographic instruction. *Research Strategies*, 3(3), 116-124.
- Torok, A., & Hurych, J. (1986). End user online searching among university faculty. In J. M. Hurd (Ed.), *ASIS '86* (Proceedings of the 49th annual meeting of the American Society for Information Science, Chicago, Illinois, 28 September-2 October, 1986) (pp. 335-40). Medford, NJ: Learned Information.
- Trivison, D.; Chamis, A. Y.; Saracevic, T.; & Kantor, P. (1986). Effectiveness and efficiency of searchers in online searching: Preliminary results from a study of information seeking and retrieving. In J. M. Hurd (Ed.), *ASIS '86* (Proceedings of the 49th annual meeting of the American Society for Information Science, Chicago, Illinois, 28 September-2 October 1986) (pp. 341-49). Medford, NJ: Learned Information.
- van Styvendaele, B. J. H. (1977). University scientists as seekers of information: Sources of reference to periodical literature. *Journal of Librarianship*, 9(4), 270-277.
- Vincent, C. P. (1984). Bibliographic instruction in the humanities: The need to stress imagination. *Research Strategies*, 2(4), 179-184.
- Weinberg, B. H. (1987). Why indexing fails the researcher. In C. Chen (Ed.), *Information: The transformation of society* (Proceedings of the 50th annual meeting of the American Society for Information Science, Boston, 4-8 October 1987) (pp. 241-44). Medford, NJ: Learned Information.

- Weintraub, K. J. (1980). The humanistic scholar and the library. *Library Quarterly*, 50(1), 22-39.
- Wiberley, S. E., Jr., & Jones, W. G. (1989). Patterns of information seeking in the humanities. *College & Research Libraries*, 50(6), 638-645.
- Wilson, P. (1977). *Public knowledge, private ignorance: Toward a library and information policy*. Westport, CT: Greenwood Press.
- Wilson, P. (1980). Limits to the growth of knowledge: The case of the social and behavioral sciences. *Library Quarterly*, 50(1), 4-21.
- Wilson, T. D. (1981). On user studies and information needs. *Journal of Documentation*, 37(1), 3-15.
- Wood, D. N., & Bower, C. A. (1969). The use of social science periodical literature. *Journal of Documentation*, 25(2), 108-122.