Conceptual Organization and Retrieval of Text by Historians: The Role of Memory and Metaphor

Donald Owen Case
Graduate School of Library and Information Science, University of California, Los Angeles, Los Angeles, CA 90024

As electronic text files increase in number and diversity, the problem of devising a more effective information retrieval interface grows more important. Future designs may draw upon cognitive theories of categorization and metaphor to understand how users interact with text—both paper and electronic. Relevant literature in cognitive psychology and information science suggests the importance of the user's physical environment in thinking about abstract entities, such as categories of documents. Empirical studies have established a basis for understanding how we think about, file, remember, and locate text. Results from a study of 20 historians—an exemplary group in terms of their close studies of texts and the broad scope of their inquiries—imply the importance of metaphors for storage and retrieval of documents. The study found that metaphors and subjective categories were frequently applied to documents collected and created by these scholars. Two physical factors—spatial configuration and document form—were often considered before topic in determining document storage locations in the office. Developers of information systems should consider qualitative aspects of cognition in their designs. System developers might also consider segmenting the audience for computer interfaces, as well as designing generic tools that apply to all users.

Introduction

People who work with printed information may need to keep track of hundreds, or even thousands, of paper and electronic files. While storing and retrieving paper documents efficiently has always been a challenge, the growth of electronic text files has made the situation increasingly more problematic. Not only must individuals manage files on paper, but they must also keep track of documents that may never reach paper, e.g., word-processing files.

Received February 1, 1990; revised September 8, 1990; accepted October 1, 1990.

Jones (1986) notes that most computer systems use primarily name-oriented approaches in their depiction and display of files, and that the growth in named files has led to problems of *recall* ("what did I call that file?") and *recognition* ("what's in that file?"). Thus, one of the major problems facing system designers lies in providing methods for naming, organizing, indexing, and displaying files in such a way that users can effectively retrieve the right ones. We do not fully understand how text files are best depicted on computer screens so that they match the user's mental model of how files are stored and retrieved (Borgman, 1986). Partly this is because we do not fully understand how individuals name, file, and find paper documents. Thus, the problem of filing and retrieving documents—whether paper or electronic—is of particular relevance to designers of the computer interface.

Information scientists have long studied user interaction with both printed and computer files. This article suggests how future models of text storage and retrieval can be informed by research on subjective aspects of cognition. One source of models are the experiences of actual users, as collected through qualitative interviews and inventories. This article begins by reviewing cognitive research of possible relevance to interface design, and then applies those ideas to a study of the text-handling of American historians.

Relevant Research in Cognitive Psychology and Information Science

Use of Categories and Metaphors

The field of cognitive psychology has contributed greatly to our knowledge of human information-processing. *Categorization* is a topic of particular relevance to the design of information retrieval systems. The tendency of humans to sort their experience into labelled categories has been addressed by such scholars as Ludwig Wittgenstein (1953) who developed ideas of

extendable boundaries and family resemblances, and Lofte Zadeh (1963) who discussed gradiated membership in "fuzzy" sets. Categorization is thought to be fundamental to human thinking: "virtually all cognitive activity involves and is dependent on the process of categorizing" (Bruner, Goodnow, & Austin, 1956, p. 246).

Recently the researcher most closely associated with categorization theories has been Eleanor Rosch (1975, 1978), of the University of California at Berkeley. Rosch was the first to develop systematic explanations of the effects of prototypes in thinking. Her work, and the general importance of categorization to human thinking, were featured in George Lakoff's *Women, Fire and Dangerous Things* (1987), which introduced this aspect of cognition to a larger audience. Rosch and her colleagues have an impressive record of published studies that establish the principles of natural categorization and connect it with other research topics in cognition, such as mental imagery and semantic representation.

Based on work by Rosch and others, Johnson (1987) and Lakoff and Johnson (1980) have offered compelling evidence that much of our language-related behavior, such as categorization, is grounded in our physical experience. Studies of language usage show that thought is *embodied* (that is, it is grounded in bodily perception, movement and physical experience) and *imaginative* (i.e., it uses metaphor, imagery and other imaginative relationships between ideas and things). The strongest evidence for this view comes from language behavior, in which many of our expressions indicate that we understand and experience one kind of thing in terms of another—that is, our thinking is often metaphorical. In particular, Lakoff and Johnson (1980, pp. 17–18) point out that most fundamental concepts are organized in terms of spatial metaphors: "In some cases spatialization is so essential a part of a concept that it is difficult for us to imagine any alternative metaphor that might structure the concept. In our society 'high status' is such a concept."

The relevance of cognitive theories have been recognized in artificial intelligence, which has used them to model human information processing (for examples see Rich (1979) and Gentner and Stevens (1983)). Within the field of information science, various authors (Najarian, 1981; Evans, 1982; Ingerwersen, 1982; Fine, 1984; Ellis, 1989) have also suggested that we need to delve deeper into human cognition in order to improve both manual and automated information systems. Neill (1990) has specifically identified the body's influence on thinking as a difficult hurdle for the representation and retrieval of knowledge.

Thus far, the research reviewed here implies that computer inferfaces that will be most effective are those that correspond most closely to the physical environment in which we deal with documents. These principles have in fact been applied to the computer interface and are popularly known as the "office" or "desktop metaphor" for computer storage. Nonetheless, the appropriateness of the office metaphor—and the use of analogies in general—has been challenged by various researchers. Halasz and Moran (1982) view analogies of complex systems—such as computer procedures and actions—as potentially limiting, misleading, or otherwise harmful to user understanding. The most compelling evidence against the use of analogy in the computer interface has been assembled by Susan Dumais of Bellcore, who has conducted a series of experiments (Dumais & Jones, 1985; Jones & Dumais, 1986; Dumais & Wright, 1986) comparing various ways of filing computer objects. While they have demonstrated that physical analogies do less well than simple naming, the experimental tasks do not accurately reflect real world environments or behaviors. As Kwasnik (1989, p. 147) says, it is impossible to validly conduct research on labeling of meaningful documents outside of the individual's usual environment.

Other researchers are less condemning of analogical applications. Guastello and Traut (1989) find that "mixed modality metaphors—icons that included three to eight character names as well as pictorial representations—are more effective than either names or pictures alone in the design of the computer interface. Grudin (1989) agrees with critics of analogical approaches to the user interface, noting that correspondences to the real world in user interface analogies *must* break down at some point. Yet the many counter-intuitive examples he presents show that *any* strict consistency in user interfaces can be limiting, whether based on analogy or not. Halasz and Moran (1982) point out that analogy can be effective in portraying specific points of a complex concept.

It is clear that the effective use of analogy and metaphor must be based on knowledge of the potential users—not all icons or actions will be familiar to everybody. Guastello and Traut (1989) show that population characteristics are related to the ability to recognize interface icons, and can be used to improve their design. Restorick (1986) has suggested the importance, in filing system designs, of matching the user's perception of the way information is characterized and stored in conventional office settings. Hence, the design of information systems and interfaces is best based on studies of user characteristics, as also suggested by Swift, Winn & Bramer (1979), Borgman, Case, and Meadow (1985) and Croft and Thompson (1987).

How can we learn more about human cognition in regards to the categorization, storage, and retrieval of documents? Through examinations of the physical settings in which users work with documents and carry out storage and retrieval actions. Many scholars of information have focused their attention on library settings (a public arena for information). Yet other studies have approached this topic through examinations of the personal office (a private knowledge environment). This

article follows the latter approach in studying personal information systems.

Studies of Document Categorization

Anyone who works with artifacts of information develops ways to organize them for efficient retrieval. Among the items that an office worker might need to organize are publications (journals, books, individual articles, and papers), correspondence, source materials and notes, and nonprint media (magnetic computer tapes and diskettes, audio and video tapes, optical disks, and photographs, illustrations, and other forms of visual expression). Somehow individuals organize personal collections of files to optimize their own patterns of use, yet also in ways that enable them to relate their files to the work of others. Cognitive categories, while individualized, are also somewhat consensual in nature: they are shaped by agreement within a social group, such as colleagues in the same research field (Swift, Winn, & Bramer, 1979; Douglas, 1986). For "knowledge workers," organizing the information they use must represent an enormous chore. How do they do it effectively?

Until recently, few investigators have examined how people actually organize and store information important to them. Among the few early examples are Jahoda, Hutchens, and Galford (1966) and Soper (1976). Their studies used survey techniques to examine personal collections and indices, focusing on *what* people did with documents and *how* they did it, but not *why* they chose to do it in the ways they did. A similar pattern is followed by texts such as Jahoda (1970) and Stibic (1980), which concentrate on *how* to create optimal personal files.

While information is usually captured in physical form, its labels and meaning are often culturally-determined, as suggested in research on information retrieval by Neill (1983). We need to understand users' motivations in document storage and retrieval, how they impose meaning on text, how they develop labels for categories of documents and how they retrieve them at later times.

One of the first attempts to understand the "why" of personal information systems was the study by Cole (1982). Cole surveyed 30 office workers and their filing systems, concluding that the six most important aspects of information were "type, form, volume, complexity, functions, and levels of storage," although the first five of these were not fully defined by him. Regarding the sixth aspect, Cole said his respondents interacted with three "levels" of information in their files: "action information" (documents readily at hand and sometimes found in stacks on office surfaces), "personal work files" (held in conventional filing devices and furniture) and "archive storage" (information stored away from the office). As information moves across these three levels, from "action" to "archive," the user's understanding of the spatial element becomes less vital while knowledge of the category structures becomes more important. Frequent interaction with the "action" information is necessary to prevent the spatial clues from becoming indistinct. In the case of personal work files, Cole noted that retrieval can be based on memory of a file's location or by 'mapping' storage schemes onto physical locations. Archives are dealt with only rarely and almost exclusively through an extensive category structure. Cole called for further research to investigate the best ways to implement spatial dimensions in the computer interface and to enrich temporal (e.g., date) and physical (e.g., color, shape, and size) retrieval cues for electronic files.

Several other studies have examined people who work with literature of various types. Malone's (1983) case studies of research scientists and clerical office workers were sparked by his interest in designing better computer interfaces. He made three primary observations about the way these information workers organized their information. First, Malone found that the difficulty of categorizing information was the critical factor in determining how people arrange their working materials. Respondents organized much of their working materials into small stacks around their offices according to "type," such as articles and notes relevant to a particular project. But the typical respondent's office contained more loosely defined piles of mixed content, many of them arranged chronologically, in which in the order of placement in the stack serves as a cue to retrieval. A second finding of Malone's was the tendency of respondents to arrange materials in such a way as to remind themselves of what to do, as well as to help locate information. Thus, there were piles of correspondence to be answered, journals to be read, and general lists of things to do. Malone found that the proximity of these stacks to the respondent's chair was an indicator of the importance or urgency of the tasks it represented: the closer it is, the greater the need of attention—a metaphorical use of space. Finally, Malone identified two common types of document processing problems: items that defy ready action or categorization, and items that one needs readily at hand. Both often end up in piles, the number of which are related to the worker's style and type of work (e.g., routine or nonroutine).

Case (1986) interviewed 60 university professors in their offices and observed the number of stacks of materials lying around the office; the linear footage of books, journals, and notebooks on the shelves; and the number of filing cabinet drawers and other storage devices. Despite no obvious shortage of other storage devices, he found that respondents kept an average of 19 (and as many as 49) stacks of documents in their offices. Thus, the variety of workers studied by both Malone and Case used a great many stacks as a physical scheme of organization, suggestive of a backlog of docu-

ments that were either frequently used or difficult to classify.

Kwasnik (1989) used the methods of cognitive anthropology in her study of the influence of context on the way in which eight professors classified documents in their offices. Her qualitative study suggested that *context*—the situation, predisposition or history of the individual—sometimes plays a more powerful role in categorization decisions than the attributes of the document itself. Categorization of one thing always takes place in relation to other things—a prototype of a category, for example. The descriptive dimensions that Kwasnik found to be most often used by subjects in their classification decisions included *form* (e.g., "book"), *use* or *purpose* (e.g., "to give a lecture..."), *time* (e.g., frequency of use of, or age of, a document), *topic* (e.g., "role reversal"), *location* (e.g., "things on the desk") and *circumstance* (e.g., an ongoing project). Kwasnik's investigation provides an in-depth picture of the way in which people identify important concepts, organize meanings and form rules for categorization.

All of these studies suggest a tradeoff between physical and mental efforts in the categorization of text. That is, that sometimes we use only spatial orientation to organize documents and at other times we use space in combination with explicit labels for document categories. The investigations by Cole, Malone, Case, and Kwasnik also suggest both similarities and differences among various user populations. Since the studies were not parallel in methods, it is difficult to tell how similar the academics studied by Case and Kwasnik are to the office workers in the Cole and Malone studies.

A reading of this literature suggested a further study of filing in the office. Research questions, results, and their interpretation follow.

Research Questions

Among the research questions suggested by earlier investigations are the following:

- Do the same facets of storage noted by Cole and Kwasnik appear in scholarly work, e.g., the writing of books and articles on history?
- Do scholars think about the contents of documents as having a physical form, e.g., do they talk metaphorically about the concepts they discuss in their written work?

These questions, coupled with a desire to learn more about the nature of historical scholarship, led the author to conduct the study described below.

An Example of Text Organization and Retrieval by Historians

Dimensions of the Investigation

The suggestions and insights presented here were gathered during a study of the work of American historians, undertaken to better understand scholarly work in general and that particular community of scholars. This was a qualitative study of scholars in the setting of their offices. In line with the advice of Glaser and Strauss (1967) and Turner (1981), the study included overlapping data from various sources, in order to build a thicker description of the phenomena under study. The sources of data included lengthy interviews, observations taken in the workplace and at conferences of historians, and examinations of documents written by respondents and other historians.

The study examined documents collected and abstracted from outside sources as well as text produced by the scholar. Investigated was the labeling and storing of relevant passages of text, and the experience of searching through those items. Observations focused on documents collected by the historian, "abstracting" of those documents (by way of glosses, notes, and indexes), categorizing of the text and the creation of new documents through the historian's writing process.

The Historian as Expert User of Text

Historians were chosen as subjects for a variety of reasons, not the least of which is that there have been precious few studies of these scholars actually *do* in their work (notable exceptions are Uva, 1977, Stieg, 1981, and Orbach, 1984). According to one scholar (Tilly, 1981, p. 14), historians are people who read, condense, collect, assimilate, transform, and synthesize written records of past times and commentaries on those records. The primary task of historian lies in scanning a textual environment for stimuli (books, periodicals, letters, diaries, etc.) that match certain characteristics, and absorbing and interpreting those stimuli in terms of a larger historical theme or theory.

Germane to this article, historians are a fruitful group to study because, in order to do their work effectively and efficiently, *they must develop techniques to identify, scan, index, file and retrieve large amounts of written material relevant to their projects.* Compared to many other academic specialists, the work of the historians may involve a broader range of methods and topics, a circumstance that would allow us to relate their information behavior with that of knowledge workers in several other fields. Although their work encompasses some of the same topics studied by the social sciences, it emphasizes the dimension of *time*—often ignored by other disciplines that study human behavior.

While the historian's tasks are largely cognitive, it is the interaction between their mental and physical processing of written material that is at the heart of this study: how is it that the texts selected by the historian as "relevant" are indexed and stored for later retrieval? That is, how does the historian manage—through spatial placement and physical and mental labelling—the large quantities of text that are gathered in the process of creating an "original" work?

Obviously the work of historians may vary with the stages of their research projects. Connecting all phases, however, is the need to "collect" (in at least a mental, and more often a physical sense) portions of written information, and to "store" them for future use. The nature of historians' work makes them a useful example of the effective use of categorization in a cognitive task—a population that could provide us with insights on the indexing, storage, and retrieval of documents. Among other things, historians are experts in managing text.

Methodology

Over the winter of 1988 preliminary work was accomplished, including literature reviews and development of an interview schedule and observation methods through pretest interviews with two historians. Once the instruments and procedures were ready, intensive interviews, typically lasting 60 to 90 minutes, were conducted with an additional 18 historians over a five-month period during the spring and summer of 1988.

Respondents were selected from eight universities (two private, six public) in the states of California and Arizona. Four of these universities were "research" institutions, as so designated by the Carnegie Foundation (1987), and accounted for 12 respondents. The other four institutions were primarily devoted to teaching and included one "doctoral," one "comprehensive," and one "liberal arts" institution according to the Carnegie classification; together, the latter four universities provided eight respondents.

For purposes of comparability of interests (in terms of geography and time periods), the sample was restricted to historians of the United States. The sampling frame used were lists of full-time faculty at each university, taken from the *Guide to Departments of History,* 1987–88 (Vincent-Morgan, 1987). Faculty identified by their departments as "on leave" were not contacted. A total of 94 potential subjects identified on these lists were contacted by telephone and/or visits to their departments. Nearly all historians successfully contacted were subsequently interviewed (there were four who declined or postponed an interview). Thus, an additional 18 historians were interviewed, bringing the total to 20 (21% of the sampling frame). The sample is not assumed to fully represent the 16,000 or so American historians resident at United States universities.

The sample is biased in favor of males (19), more senior faculty (11 professors, three associates and six assistant professors) and probably in favor of faculty who spent more time in their offices. By and large these were accomplished scholars: most (15) had written at least one book and one as many as seven; most had also published several articles and one more than 20.

Verbal responses were taped as well as recorded on an interview schedule. Later the two types of records were merged to form selective transcripts, organized by interview topic. Questions and observations covered research and teaching interests, current research projects, methodologies used in research, nature and location of materials used in research, personal filing and indexing methods (for books, articles, notes, etc.), numbers of shelves for books, journals and other materials, numbers of filing drawers (and their contents), equipment and furniture in the office, and use of computers and other technologies. Regarding the historian's files, the distribution of stacks of papers and books was observed and questions were asked about their content and purpose. The interviews discussed the terms created for labels of subject-oriented files, the degree to which time-oriented categories were used, and what kinds of items are difficult to categorize and/or retrieve from storage.

Sketches were also made of the historians workplace. These were used to analyze the ways in which the respondents dealt with information in its physical form. Citations to the published work of the repondents also were collected in order to better understand the topics and methods of each respondent through an examination of sample articles and books. These were complemented by examination of historical textbooks and published descriptions of historical methods by noted historians—material too broad to be covered here (see Case, 1991).

Results

The following passages describe patterns noticed across several (but rarely all) interviews, and make use of quotations and examples. To provide the reader with a better idea of how many historians involved in a particular generalization, the interviews from which the conclusions or quotations are drawn are noted by the interview numbers (1–20, in parentheses).

The Office Scenario

With three exceptions, offices were crowded with furniture, equipment and printed material. The exceptions (10, 12, 14) were respondents who had very small office spaces, including one who was forced to share an office with another instructor. Objects commonly found in historians' offices were:

- Furniture (desk, bookcases, file cabinets, chairs, etc.)
- Equipment (phones, answering machines, computers, microfilm readers, etc.)
- Personal belongings (clothing, dishes, food, radios, art objects, toys, etc.)
- Artifacts of information: books (personal and library copies), journals (personal and library copies), newspapers, articles (individual copies—photocopies and reprints), notes (taken by the historian), manuscripts (of work in progress), magnetic tapes (computer, audio, and video), and visual materials (photographs, slides, prints, etc.)

The scholars interviewed tended to keep a great deal of printed and other information in their offices: books and journals on shelves, papers in and on top of filing cabinets, and papers, books and journals spread out over desks and tables. The sheer amount of these materials is difficult to estimate, given the variance among factors such as length and "fullness" of shelves and file drawers. For example, it can be said that the number of shelves of books ranged from as few as three to as many as 64, but the shelves themselves ranged in length from 28 inches to 42 inches, and were found in varying degrees of fullness. Working habits also varied widely, with some respondents doing virtually all of their work at home—and so having few stacks of paper in their offices, while some had dozens of piles of books, articles, and papers around their offices. Suffice to say that the model respondent had in his or her office 35 full shelves of books and journals, two four-drawer filing cabinets, and 12 stacks of materials representing some kind of teaching or research activity in progress.

It is worth noting that even scholars who work primarily at home use their university office as the main storage area for the teaching and administrative aspects of their work. Thus, studying the secondary workplace may provide some clues as to how their main workplace is organized.

Organization by Spatial Constraints

As in most offices and libraries, the primary object was to keep like things together while satisfying other constraints. These other factors included the need to keep some things close at hand (e.g., frequently used material), to fit within the particular physical confines of the space (e.g., available bookshelf space, space and size of room) and to provide some kind of reminding function (e.g., keeping materials for this week's classes on the floor, where they would be impossible to overlook).

"Keeping like things together" can mean a number of things. As in libraries, scholars also strive to keep documents on similar topics as close to one another as possible, within the limits of three-dimensional space. Yet above that all-important level of organization is suspended another: it is often desirable to keep like *forms* together. Obviously some forms simply do not go well together physically: it is difficult to intersperse individual articles with books, for example—although two respondents did this by using paperboard containers for groups of papers.

Yet some forms easy to interfile were not found together. For example, respondents usually keep journals on separate shelves (or even sides of the room) and sometimes store hard-bound books separate from paperbound books. In terms of subject classification, there is no justification for doing the latter, but two other reasons for doing so appeared in the interviews. The first is aesthetics—to some, if books are not of about the same height and construction they simply *look* wrong when placed together—a factor that this discussion will ignore as too much a matter of individual taste, as with the respondent (19) who kept "old books, or at least books that appear old" on a separate shelf. But the second reason is relevant to the discussion: the shape and feel and look of the book were a part of the experience in both reading and subsequent recall of the material. One would not look at the shelves of paperbound books to find the copy of Carr's *What is History*? because one could remember that it was a hardbound book. And one would not bother to check the index for the passage about Karl Popper, because one knew that it appeared three-fifths of the way through, on those pages with the most margin notes.

It might seem an obvious point that since we have to live in a physical environment, that form would take precedence over content in some situations. After all, libraries emphasize form over content when they place oversized or undersized books together—purely an accommodation to the efficient use of shelves and racks. But it is important to recognize that physicality is important even to those whose primary job is to emphasize content. Whatever their degree of concern with ideas, historians also consider the form in which they are written.

At the top level of organization, then, the factors were largely physical: consideration of spatial constraints, followed by attention to form.

Organization by Form

It was with books that the broadest categories (and sometimes subcategories) were evident in the descriptions which respondents gave of their groupings. In contrast, journals were typically kept separate from books and grouped simply by title.

While a few respondents went to considerable length to index their "holdings," indexes to books and journal articles were usually simple, if they existed at all. As revealed by respondents, much of the "organization and retrieval" of texts occurs during the act of writing and not before. The function of many of the cards and computerized lists used by historians were to capture and sequence ideas and references to be used in writing a specific document. Many of the cards did not exist before production of the document was begun, although they were often maintained after the writing project was over.

In the most elaborate example of a system to aid writing, a historian (12) of black Americans maintained hundreds of thousands of 5 × 7-inch cards, held with rubber bands in metal card files. These cards were used to create detailed chronologies of what the most important subjects in his books had done during their lives. These were organized first by year and month, then by personal name, then by specific event (e.g., "Harper's Ferry") or continuing events (e.g., "Social Life"). Cards

were also used to record relevant bibliographic references, facts and problems to think about. Some cards were added or updated virtually every day.

Twelve (60%) of the respondents maintained some kind of card file to index the books and/or articles they had collected and read in support of their scholarship. But very few of these were like library catalogs in either their purpose or exhaustiveness. In some cases, the indexes did not attempt to inventory everything the scholar owned, but rather only those relevant to a writing project currently underway—three were keeping such a card index at the time of the interviews. Several (13, 15, 17) had tried to develop an index for their entire document collection, but had abandoned it. One respondent (13) had tried with a controlled list of terms as many as five times, but gradually gave up each attempt. In this respect the historians seemed similar to the scientists described by Burton (1981): sporadic efforts to build or update an indexing/filing system that is not otherwise maintained.

Despite an antitechnology bias in a tradition-oriented profession, it is clear that many historians are willing to try any device that promises to help them with their filing and writing. Many of them had already tried various manual or automated systems for keeping track of their literature, often abandoning it after a time. In this respect they were quite like the individuals studied by Cole, who concluded that (1982, p. 60): "Individuals often start to construct an elaborate filing system, including the use of color coding and hierarchical classification, but discontinue it after a time." Dumais (1988) notes that indexing in the paper-and-pencil environment entails high costs for the benefits received in later retrieval, whereas in a computer system the trade-offs are more favorable. Historians in this study were sometimes skeptical about the prospect of computers or even manual tools helping them with their work (11, 15, 17). One respondent (15) reached back 20 years in time to ridicule the use of rod-and-card tickler files of cards by himself and his colleagues, before continuing on to scoff at the present-day obsession with computers.

Organization by Topic

It is only at the third level of organization that we reach what we typically think of when we discuss intellectual access: "topic" or "subject." Respondents were asked to go "around" their office and talk about what was there. A typical statement came from respondent #9, whose entire book collection was in his office, organized by topic

> and somewhat chronologically... over there you have some things on American literature and most of these middle files concern legal history... textbooks, general textbooks on American history... and then I go chronologically: Colonial Period, Late 19th century, Civil War, Reconstruction, Progressive Era, New Deal, Cold War, Foreign policy... yeah, they are sort of chronological through sort of a mixture of chronological and topical.

Respondent #19 had the following categories (each a separate shelf, except for three categories that occupied two shelves each)

> Western civilization, American literature, books about American literature, Old books... duplicates of books at home... general religion, sociology, twentieth century, women's studies, black studies, textbooks, miscellaneous books, writings of Thomas Jefferson, English History... In theory the books on each shelf have something in common.

Several respondents kept most of their books at home, or used few to begin with, but still used a two-level scheme in which books were organized topically within chronological periods.

Most historians had faith and even pride in their categories. Respondent 11 says of his categorization scheme: "By this time I know what I'm doing... my course, my books, and my files are organized about the same way, a combination of periods and topics and I find that it works very well.... most historians don't even think in terms of subject authority... it's [only] the sort of thing that you would have to do if you used a computer."

Some used geographical designations, such as "the Western movement" or "Latin America." At times the labels for chronological periods were those recently adopted by the *Journal of American History* (Gilmore, 1989), which seems to be responding to their common use among historians. One respondent (20) described the value of chronological organization in his work in these words:

> What is difficult is if you don't want to keep making up a lot cards for each one... the reason that I always start by putting it in chronology is I'm not necessarily sure of what the categories until I see the lay of the land is on the project itself. So I start off by saying that I want to do something on labor and the movies. But what exactly? I'm not absolutely sure... So the easiest thing to do when you start off, when you're looking through microfilms and manuscripts is to put things chronologically... Then I just go through all the cards and create the categories as I'm going along.

Most (15) respondents chose to organize their files and notes mainly by some arrangement of topics. Among those 15 historians, 10 chose to organize also by chronological period or event, either in combination with topical categories or as a secondary level. Two historians said they organized their material strictly chronologically, and another one by topic within periods. The remaining two historians claimed to organize their materials only by author's name—which could be considered a kind of indirect subject organization.

Organization by Treatment, Purpose, or Quality

The next level of storage was a more specific level that defies a comprehensive label. At this level begins to emerge some distinctions among broad topics that are not necessarily hierarchical. For instance, one group of materials might be of a certain intellectual genre (e.g., biographies). One respondent (19) even broke out a shelf by *quality* ("'Good' recent American intellectual histories, 'bad' recent intellectual histories"). Many respondents kept textbooks separate, either within or outside of broader topic areas—a distinction by purpose.

In summary, apparent in virtually every interview was a concern with four levels of storage:

- *Space* (physical space, as configured by walls, bookshelves, etc.)
- *Form* (single page, multipage, paperbound book, hardbound book, other media)
- *Topic* (broad subject, e.g., "foreign affairs," or a time period, e.g., "early national")
- *Treatment* (e.g., biographies) or *purpose* (textbooks, books for class) or *quality* ("good" versus "bad" examples).

Here a parallel appears with cross-cultural studies of color perception and naming, which show that black and white are the first colors to be university recognized with a name, followed by red and then either yellow or blue or green. A similar pattern emerges in offices: it is the physical space that is first given priority in document location, and then the physical form of the document. It is the third level that contains the factor typically of most concern to the information scientist: that of the specific topics of the document. (Topical *relevance,* it should be noted, it is implied by the document's very presence in the office.)

Forming and Labeling Categories

Because of the variety of research interests represented in the sample, it is difficult to compare results across historians. Here, selected quotations may convey the flavor of the respondents' thinking. Historian 17, for example, described his process of organizing materials in this way:

> I've tried to do it by subject and that's very difficult because you start out with maybe a sort of very rough hazy approach... and you find out in a little while that all those subjects are irrelevant, so you reorganize, that's the problem... You get a lot of notes and you start losing things. What subject did I put that under? Where is that thing? When you're trying to develop a new idea, when you're trying to get a handle on it, concretize it, systemize it, find out what it really is you're talking about: it's awfully hard to put a label on it. It will usually be about a year later, when I'm lecturing, trying to explain it to students, that some label will pop up.

Yet another respondent (20) emphasized how closely intertwined is categorization with his writing:

> When I get ready to write, I go through the cards and I reorganize them into categories. For example, I will pull cards that have to do with specific films made by labor. I will pull cards that have to do with specific government agencies that got involved with filmmaking, and put them by agency. I will pull cards and put them in a category "Censorship." In other words, [for] the main categories that I am going to write on... once I see what it looks like its easy for me to go through the cards and put them in what I can see are going to be the sections of either a book or an article.

One respondent (12) expressed the comments of others when he described his problems in keeping track of things by the labels he had assigned to them: "So often a piece of material falls into two or three categories. You can't constantly take it out of one and put it in the other files, but you can make duplicates of it and the computer is going to be of value in doing that kind of thing." In a similar vein, another historian (19) describes why he sometimes files duplicates in one or two additional folders:

> In a more slipshod system you lose track of stuff and the critical moment comes and you can't find it when you need it. And you waste lots of time looking for it. I tend to have that problem, practically every day! The problem is that what I file could be filed in any one of several folders... I don't always remember what the decision was.

Like the students studied by Kuhlthau (1988), these respondents would sometimes express anxiety when discussing the early stages of a research project. More experienced scholars learn to conquer their fear of uncertainty when approaching new material. Presumably this is because accomplished scholars have learned how to adroitly develop categories for their new research and "saturate" them with examples that clarify their boundaries. From this stage follow definitions of phenomena illustrating the categories, recognition and exploitation of links between them, and eventually, hypotheses and theories concerning the phenomena under study.

The Role of Memory

Historians claimed to rely much less on topical organization of materials than they do on their own memories for where a theme or particular passage can be found. Although they differed widely in the degree to which they organize the physical artifacts of knowledge, most historians in this sample were alike in claiming to have an excellent memory of the location and meaning of various items of text. As respondent 11 claimed: "I know where everything is." Historian 20 emphasizes that "I know my notes. I read through them

every few months, read through the whole thing." Respondent 8 declared that "I would have to say that I rely a lot on my memory... of a particular interview or of particular documents I've read, and that I trust in my ability to recall those. I wouldn't say that I have a photographic memory but I am able to have pretty accurate recall of a lot of things."

Some (12, 18, 19) admitted problems in finding certain documents: respondent 17 described his retrieval efforts as consisting of "memory and hard work. I can spend hours locating something I've got somewhere." When historian 8's memory failed it was usually when he would "recall a particular document having a certain content that it did not," rather than not being able to find the document. Respondent 12 complained that he "loses notes all the time... I haven't yet figured out a filing system that is useful" but maintains that he has "a fine memory" something any good historian must have. "Often I remember what obscure file contains this little piece of evidence."

This ability of historians to remember so many details should come as no surprise. Since at least the time of Simonides, we have known that it is possible to memorize many names or facts based on visualization of them in physical structures. The Jesuit Matteo Ricci was able to recall thousands of items of information by imagining "memory palaces" with hundreds of rooms—each equipped with unique furniture and images (Spence, 1984).

The very first respondent in this study (1) warned that historians' offices would reflect "an untidy world." The respondents were, in fact, frequently apologetic about their disorderly offices (11):

> this mess that you see here that could be relieved by another filing cabinet is essentially a file which includes everything from newspaper magazine clippings to manuscripts to transcripts of oral history to interviews to article copies.

Dwyer and Dwyer (1987) point out that the effectiveness of retrieval cues in human memory is dependent upon the elaborateness of the encoding process during learning. That is: what the learner does during learning is intimately related to the retrieval of the information absorbed. They also point out that the environment in which learning takes place is vital to subsequent retrieval: any changes in the environment lead to decreases in performance. One of the worst things that could happen to a scholar would be for a stranger to "tidy up" their office. The act of arranging things in a "neater" fashion would destroy many important connections, patterns, and associations.

Crowding does not necessarily hinder cognitive recall—even though it may have an effect on physical retrieval—since recall and recognition are different cognitive processes. Since organizing materials is difficult, one relies upon recognition as much as possible. Part of what makes an intellectual workplace "effective" is its very crowdedness: the goal of having important and frequently used things in view and close at hand. This in turn inevitably results in messiness and social if not cognitive discomfort. The presence of humor often indicates the taboo, the truth that cannot be spoken: our jokes about untidy desks turn on a common belief that a messy office implies a disorganized mind—a metaphorical connection that this investigation suggests is untrue.

Through repeated strengthening of their mental associations between documents and concepts, historians are able to conduct successful searches in even the most disorganized files. This even extended into visualization of the particular page or portion of a page, on which a passage appeared. Rothkopf (1971) has demonstrated that readers may picture the position of text on pages, relying upon visual and tactile cues to aid recall.

Use of Metaphor

Metaphorical references that compared ideas to physical entities were encountered in nearly every interview (italics are the author's)

> "questions and ideas that are *swimming* in my head" (ideas as fish)
> "I *cast my net* very widely" (ideas as fish)
> "*sitting down*" and figuring it all out (ideas as puzzles)
> "*cut and paste*" thoughts together (ideas as pieces of paper)
> "keeping *track* of things" (document storage as game hunting)
> "trying to *get a handle on*" and idea (ideas as tools)
> "to *put a label on*" an idea (ideas as jars)
> "we *borrow and steal* from each other constantly" (ideas as valuables)
> "I *toyed* with" ideas on cards (ideas as playthings)
> "*wrestle* your way through the literature" (ideas as challenging opponents)
> "*Asking questions of* the material" (ideas as people)
> "they *burrow into* their topic" (scholars as moles)
> "my interests are *bounded* by time" (topics as territories)
> "I have in my mind a *bibliography of bibliographies*" (topics as lists of relevant documents)
> "a rough, hazy *approach*" (writing as flying/landing a plane)

Most of these expressions correspond to Lakoff and Johnson's (1980, p. 47) categories of idea metaphors, such as "ideas are organisms" or "ideas are products." The ones listed above were among the metaphors that could be stated simply. Others require some explanation, as in one respondent's description of recent changes in historical philosophy as first "remodeling the house" (i.e., revising the traditional views of American history) the "tearing down the house and building a new one" (i.e., presenting a completely new view of American history). The same historian also described half-hearted attempts at revising history texts as "sprinkling powdered sugar" on the traditional views to make them more palatable to the critics. Another respondent

paraphrased neophyte historians' attitudes toward note-taking, combining two metaphors to describe how historians reach an important goal: "When you had two *shoeboxes* full of *note-cards* you should write your *dissertation*." This particular case could be considered an example of *synedoche,* in which a part stands in for the whole: a collection of ideas equals the dissertation.

As it turned out, *the card* is powerful metaphor in the work of historians. A card could represent single or multiple instances of ideas, evidence, questions, quotations, or documents. All but three of the respondents used hand-written cards or half-sheets of paper—invariably either 3 × 5, 4 × 6, or 5 × 8 inches—in keeping track of ideas and documents in their writing and teaching. The other three used only their computer to make notes and bibliographic entries. Even one scholar who had given up making new hand-written cards, still maintained his original file by using his computer to print out his notes as 5 × 8 pieces of paper, merging them with his older cards. In at least four cases, the number of cards maintained in this way were in the thousands. Historians are often trained to collect information in card-like chunks. Guides to research methods (e.g., Furay & Salvevouris, 1988, Barzun & Graff, 1970) urge the historian-to-be to use a consistent medium and style in capturing material—typically the card.

The use of metaphor by this sample of scholars points out its usefulness in expressions of ideas at various levels of abstraction. This evidence also confirms the relevance, to scholars, of the card metaphor that has been so popular in database and hypertext systems.

Implications for Interface Design

Using Analogy and Metaphor

This study found the use of analogy and metaphor to be common among historians in their references to both ideas and documents. It is suggested that the benefits of analogical approaches may outweigh their alleged problems (e.g., Grudin, 1989; Halasz & Moran, 1981; Guastello & Traut, 1989) and could be applied to information systems in special domains, such as history. System metaphors may help users remember systems actions by embedding functions parallel to those in their working milieu.

If thoughts and ideas are embodied, then it makes sense to depict them visually and spatially on the user interface. "Use" then becomes a movement through knowledge spaces in search of a physical location for storage. In this sense the film *Tron* (1982) may have had the right idea in depicting computer memory as having a world of its' own.

Visual Display and Manipulation

While the vocabulary common to American historians could be a target for implementation (see Swift, Winn, and Bramer (1979) for an example of the social scientist's vocabulary), the use and relative importance of terms varied widely among the historians interviewed, except in the areas of time and geography. Ways to extend the office metaphor for historians, therefore, could focus on chronological and geographic facets of the information, rather then standardization of the vocabulary for topics. A potential direct-manipulation device for both assigning categories and retrieving them would be a sliding scale of years, matched with labels for periods. The labels could be taken from those commonly in use by historians as well as those customized groupings devised by the individual scholar. A geographic approach could involve stylized maps of North America, with individual states ("Virginia"), geographic regions ("New England"), historical regions ("the 13 colonies") defined for easy selection. These implementations could use a mouse and visual interface for selection.

It may be too obvious to say that we need better display devices than the CRT screens commonly in use. The physical and psychological problems associated with these have received ample attention elsewhere in the literature (see, for example, Bennett, Case, Sandelin, and Smith (1984)). While other aspects of computing have changed at a bewildering rate, it is remarkable that we continue to use a display device that was invented over 60 years ago and has seen little improvement over the past two decades. It is truly a larger impediment to interface design than any software issue.

Off the screen, we need better devices for promoting not just manipulation of screen contents but tactile associations with the material. Perhaps the types of sensory gloves used in some robotics applications (in which the movement of the fingers, hands, and arms is tracked in three dimensions) can be adapted for use in document manipulation. Using the computer could eventually become more like conducting a symphony than typing on a typewriter. Isn't it a grander, more fitting image of tools for liberating the mind?

As much as is possible, the inadequacies of display and manipulation devices should not be allowed to interfere with solutions to information retrieval problems. While further tests of the effectiveness of metaphors are in order, inconclusive results must not slow innovation in interface design—particularly when the results of the experiment contradict other sources of knowledge about human abilities, including common sense. Physical analogies for procedures and information display can continue to be exploited, based on demonstrated human tendencies in information handling.

Studying Users and Systems

Regarding methodology, the study of small groups of users in depth is very time-consuming for the degree of insight it allows into human information handling, and is limited by the uncertainties of self-reports (Nisbett &

Wilson, 1977). Yet it is one of the few mechanisms we have for finding out what goes on inside people's heads. The top-down approach of qualitative studies is complementary to bottom-up approach of statistical methods of term extraction that focus on document description. A focus on terms is predicated on an objectivist view of language: the assumption that words fit reality and have fixed meanings based on strict correspondence between concepts and objects. Most of our information retrieval systems treat texts as abstract objects, in which meaning can be extracted from the parts and structure of the sentences that comprise them. Newer cognitive theories, as well as experimental research in information retrieval (Furnas, Landauer, Gomez, & Dumais, 1987) contest the view that words have fixed meanings for most people.

There are several ways of approaching interface design. Experimentalism is one path and the evaluation of system models is another. Qualitative studies of information use can also contribute by adding to our understanding of the use of metaphor in human reasoning.

Summary

The current investigation (as well as the ones conducted by Cole, Malone, and Kwasnik) were undertaken in the context of understanding how people dealt with the physical artifacts of information. Although personal filing systems and workplace arrangements are rarely discussed in the literature on information work, they must play an important role in the actual use of textual information—whether printed or electronic. Individual habits in handling the physical artifacts of information provide us with clues for the design of computer-based systems.

This study found that physical characteristics—of both documents and office—are given first priority in office filing. It is not until physical restraints are considered that document topicality is considered—the factor typically of most concern to the information scientist. This finding parallels those of Cole and Kwasnik, both of which found the form of an item to be a key determinant in storage decisions. Beyond that common element, the results of these three studies diverge along the lines of their methods and foci, with Kwasnik's being most comparable. Kwasnik's facets of location (space), topic and purpose are similar to the dimensions found to be important in the present study.

Future systems could facilitate recognition of files through the addition of richer contextual cues—i.e., such considerations as form, purpose, and quality. The user should be allowed to organize information by any cue that makes sense and at any level of abstraction. Computer systems must mimic both human thinking and the characteristics of physical storage, as well as the language of the documents. Such requirements are complementary, as psychological research indicates that cognitive processes are physically grounded in the real world. The present investigation supports that view in regards to the storage and retrieval of text.

References

Barzun, J., & Graff, H. F. (1970). *The modern researcher.* (rev. ed.). New York: Harcourt, Brace & World, Inc.

Bennett, J., Case, D., Sandelin, J., & Smith, M. (Eds.) (1984). *Visual display terminals: Usability issues and health concerns.* Englewood Cliffs, NJ: Prentice-Hall.

Borgman, C. L. (1986). The user's mental model of an information retrieval system: an experiment on a prototype online catalog. *International Journal of Man-Machine Studies, 24,* 47–64.

Borgman, C. L., Case, D. O., & Meadow, C.T. (1985). Designing an information retrieval system based on user characteristics. *ACM SIGIR Forum, 19,* 139–146.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). A study of thinking. New York: John Wiley.

Burton, H. D. (1981). FAMULUS revisited: ten years of personal information systems. *Journal of the American Society for Information Science, 32,* 440–443.

Carnegie Foundation. (1987). *A classification of institutions of higher education.* Princeton: Carnegie Foundation for the Advancement of Teaching.

Case, D. O. (1986). Collection and organization of written information by social scientists and humanists: a review and exploratory study. *Journal of Information Science, 12,* 97–104.

Case, D. O. (1991). The collection and use of information by some American historians: a study of motives and methods. *Library Quarterly, 61,* 61–82.

Cole, I. (1982). Human aspects of office filing: Implications for the electronic office. *Proceedings of the Human Factors Society 26th Annual Meeting, Santa Monica, California,* p. 59–63.

Croft, W. B., & Thompson, R. H. (1987). I3R: a new approach to the design of document retrieval systems. *Journal of the American Society for Information Science, 38,* 389–404.

Douglas, M. (1986). *How institutions think.* Syracuse: Syracuse University Press.

Dumais, S.T. & Jones, W. P. (1985, April). A comparison of symbolic and spatial filing. *Proceedings of CHI '85, Conference on Human Factors in Computing Systems,* 127–131.

Dumais, S.T., & Wright, A. L. (1986). Reference by name vs. location in a computer filing system. *Proceedings of the Human Factors Society 30th Annual Meeting,* 824–828.

Dwyer, C. A., & Dwyer, F. M. (1987). Effect of depth of information processing on students' ability to acquire and retrieve information related to different instructional objectives. *Programmed Learning and Educational Technology, 24,* 264–279.

Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation, 45,* 171–212.

Evans, N. J. (1982). Human processing of natural categories. *Proceedings of the 45th ASIS Annual Meeting, Columbus, Ohio, October 17-21, 19,* 80–83.

Fine, S. (1984). Research and the psychology of information use. *Library Trends, 32,* 441–460.

Furay, C., & Salevouris, M. J. (1988). *The methods and skills of history: a practical guide.* Arlington Heights, IL: Harlan Davidson.

Furnas, G.W., Landauer, T. K., Gomez, L. M., & Dumais, S.T. (1987). The vocabulary problem in human-system communication. *Communications of the ACM, 30,* 964–971.

Gentner, D., & Stevens, A. (1983). (Eds.) *Mental models.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Gilmore, M. B. (1989). Observations on the indexing of history: the example of the *Journal of American History. The Indexer, 16,* 159–162.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: strategies for qualitative research.* New York: Aldine.

Grudin, J. (1989). The case against user interface consistency. *Communications of the ACM, 32,* 1164–1173.

Gaustello, S. J., & Traut, M. (1989). Verbal versus pictorial representations of objects in a human-computer interface. *International Journal of Man-Machine Studies, 31,* 99–120.

Halasz, F., & Moran, T. P. (1982, March). Analogy considered harmful. *Proceedings, Conference on Human Factors in Computer Systems,* Gaithersburg, MD, 383–386.

Ingerwersen, P. (1982). Search procedures in the library—analyzed from the cognitive point of view. *Journal of Documentation, 38,* 165–191.

Jahoda, G. E., Hutchens, R. D., & Galford, R. R. (1966). Characteristics and use of personal indexes maintained by scientists and engineers in one university. *American Documentation, 17,* 71–75.

Jahoda, G. E. (1970). *Information storage and retrieval systems for individual researchers.* New York: Wiley-Interscience.

Johnson, M. (1987). *The body in the mind.* Chicago: University of Chicago Press.

Jones, W. P. (1986). On the applied use of human memory models: the memory extender personal filing systems. *International Journal of Man-Machine Studies, 25,* 191–229.

Jones, W. P., & Dumais, S.T. (1986). The spatial metaphor for user interfaces: experimental tests of reference by location versus name. *ACM Transactions of Office Information Systems, 4,* 42–63.

Kuhlthau, C. C. (1988). Developing a model of the library search process: cognitive and affective aspects. *RQ, 27,* 232–242.

Kwasnik, B. H. (1989, May). *The influence of context on classificatory behavior.* Unpublished doctoral dissertation. New Brunswick, New Jersey: School of Communication, Information and Library Studies, Rutgers University.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Lakoff, G. (1987). *Women, Fire and Dangerous Things.* Chicago: University of Chicago Press.

Malone, T.W. (1983). How do people organize their desks? Implications for the design of office information systems. *ACM Transactions on Office Information Systems, 1,* 25–32.

Miller, G. A. (1968). Psychology and information. *American Documentation, 19,* 286–289.

Najarian, S. E. (1981). Organizational factors in human memory: implications for library organization and access systems. *Library Quarterly, 51,* 269–291.

Neill, S. D. (1983). The dilemma of the subjective in information organisation and retrieval. *Journal of Documentation, 43,* 191–211.

Neill, S. D. (1990). Body English: the dilemma of the physical in the objectification of subjective knowledge structures: the role of the body in thinking. *Journal of Documentation, 46,* 1–15.

Nisbett, R. & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review, 84,* 231–255.

Orbach, B. (1984). *Historians: information needs, information seeking and the research process.* (Unpublished Specialization Paper.) Los Angeles: Graduate School of Library and Information Science, UCLA.

Restorick, F. M. (1986). Novel filing systems applicable to an automated office: a state-of-the-art study. *Information Processing & Management, 22,* 151–172.

Rich, E. (1979). User modeling via stereotypes. *Cognitive Science, 3,* 329–354.

Rosch, E. (1978). Principles of categorization. In E. Rosch and B. Lloyd, (Eds.), *Cognition and categorization.* Hillsdale, NJ: Lawrence Erlbaum, 27–48.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology, 7,* 532–547.

Rothkopf, E. Z. (1971). Incidental memory for location of information in text. *Journal of Verbal Learning and Verbal Behavior,* 608–613.

Soper, M. E. (1976). Characteristics and use of personal collections. *Library Quarterly, 46,* 397–415.

Spence, J. D. (1984). *The memory palace of Matteo Ricci.* New York: Viking.

Stibic, V. (1980). *Professional documentation for professionals: means and methods.* Amsterdam: The Netherlands: North-Holland Publishing Company.

Stieg, M. F. (1981). The information of (sic) needs of historians. *College and Research Libraries, 42,* 549–560.

Swift, D., Winn, V., & Bramer, D. (1979). A sociological approach to the design of information systems. *Journal of the American Society for Information Science, 30,* 215–223.

Tilly, C. (1981). *As sociology meets history.* New York: Academic Press.

Tron. (1982). Burbank, CA: Disney Studios.

Turner, B. A. (1981). Some practical aspects of qualitative data analysis: one way of organising the cognitive processes associated with the generation of grounded theory. *Quality and Quantity, 15,* 225–247.

Uva, P. A. (1977, March). Information-gathering habits of academic historians: report of the pilot study. Syracuse, NY: SUNY Upstate Medical Center.

Vincent-Morgan, M. (1987). *Guide to Departments of History, 1987–88.* Washington, DC: American History Association.

Wittgenstein, L. (1953). *Philosophical Investigations.* New York: Macmillan.

Zadeh, L. (1963). Fuzzy sets. *Information and Control, 8,* 338–353.