

National Library
of CanadaBibliothèque nationale
du Canada

Canadian Theses Division

Division des thèses canadiennes

Ottawa, Canada
K1A 0N4

56179

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur

Karol William Jozef Wenek

Date of Birth — Date de naissance

1 November 1946

Country of Birth — Lieu de naissance

England

Permanent Address — Résidence fixe

19 Stephen Cross Drive
Dartmouth, Nova Scotia

Title of Thesis — Titre de la thèse

Videotex Behavioural Research:
A Comparative Evaluation of Hierarchical and Keyword-Directory
Retrieval in Three Kinds of Search Task

University — Université

Saint Mary's University

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée

Master of Science (Applied Psychology)

Year this degree conferred — Année d'obtention de ce grade

1983

Name of Supervisor — Nom du directeur de thèse

Dr. Peter W.D. Dodd

Permission is hereby granted to the NATIONAL LIBRARY OF
CANADA to microfilm this thesis and to lend or sell copies of
the film.The author reserves other publication rights, and neither the
thesis nor extensive extracts from it may be printed or other-
wise reproduced without the author's written permission.L'autorisation est, par la présente, accordée à la BIBLIOTHÈ-
QUE NATIONALE DU CANADA de microfilmer cette thèse et de
prêter ou de vendre des exemplaires du film.L'auteur se réserve les autres droits de publication; ni la thèse
ni de longs extraits de celle-ci ne doivent être imprimés ou
autrement reproduits sans l'autorisation écrite de l'auteur.

Date

29 April 1983

Signature

K. Wenek

VIDEOTEX BEHAVIOURAL RESEARCH:
A COMPARATIVE EVALUATION
OF HIERARCHICAL AND KEYWORD-DIRECTORY RETRIEVAL
IN THREE KINDS OF SEARCH TASK

C

Karol W. J. Wenek

Submitted in partial fulfillment of the requirements
for the Degree of Master of Science
at Saint Mary's University
Halifax, Nova Scotia.

April 1983

Approved:

Peter F. D. Dadd

Faculty Advisor

Approved:

John W. Catana

Committee Member

Approved:

Sheldon

Committee Member

TABLE OF CONTENTS

	Page
Abstract	111
Acknowledgment	v
List of Tables	vi
List of Figures	viii
 INTRODUCTION: A RATIONALE FOR KEYWORD DIRECTORIES	 1
METHOD	29
RESULTS AND DISCUSSION	44
GENERALIZABILITY AND IMPLICATIONS OF FINDINGS	77
 References	 85
 Appendix A: Glossary of Videotex Terms	 91
Appendix B: Example Hierarchical Search	94
Appendix C: Example Keyword Search	99
Appendix D: Personal Data Questionnaire	105
Appendix E: Videotex Attitude Questionnaire	106
Appendix F: Search-Method Rating Scale	107
Appendix G: Experimenter's Instructions to Subjects	110
Appendix H: Statistical Tables	118

Abstract }

VIDEOTEX BEHAVIOURAL RESEARCH:
A COMPARATIVE EVALUATION
OF HIERARCHICAL AND KEYWORD-DIRECTORY RETRIEVAL
IN THREE KINDS OF SEARCH TASK

Karol W. J. Wenek

April 1983

The University Academic Calendar served as the experimental database in a study of videotex information retrieval behaviour. Responses on a numeric keypad operated either a hierarchical or keyword-directory retrieval system. Retrieval successes, pages accessed, search time, and time per page served as indices of search performance in single-page-solution (unipartite), multi-page-solution (multipartite), and unsolvable search problems. Relationships between performance measures and several user abilities and characteristics were examined. Pretest and post-test measures of attitudes toward hypothetical videotex applications were used to gauge the effect of exposure. Undergraduate subjects completed pre-experimental tests of spatial memory and ideational fluency, and, during the experiment, attempted 20 search problems.

Overall, the keyword directory retrieved more relevant information than the hierarchy. Unsolvable tasks required the most effort of

searchers and unipartite tasks the least. Multipartite tasks required a large number of pages and long search times but also retrieved the greatest quantity of target information. Comparisons showed, in the unipartite and multipartite tasks, that keyword searches consumed half as many pages and half as much time as the hierarchy per unit page of information. In unsolvable tasks, the hierarchy out-performed the keyword directory, taking substantially fewer pages and less time to confirm the non-existence of target information. Performance measures were found to correlate moderately with the spatial memory scores of those who used hierarchical retrieval, but with the ideational fluency scores of those who used keyword retrieval. Post-test attitude measures reflected a significant increase in favourability in the videotex application experienced by subjects. Implications of the findings and other issues raised by the results are discussed.

ACKNOWLEDGMENT

Without the benefit of Dr. Peter Dodd's considerable programming and computer skills, this experiment could not have been carried out. For the labour-saving programs which assembled the keyword directory and for the many evolved versions of the information-handling and data-collecting programs, I am especially grateful. In no lesser degree, I wish to record my appreciation of his touchstone role as a thoughtful and patient critic.

LIST OF TABLES

	Page
Table 1: Performance measures used in previous research	19
Table 2: Reliability and validity estimates for tests of spatial memory and ideational fluency	35
Table 3: Subject-group composition according to pre-measures	39
Table 4: Target pages successfully retrieved: means and standard deviations	49
Table 5: Total pages accessed: means and standard deviations	51
Table 6: Search times: means and standard deviations	53
Table 7: Time per page: means and standard deviations	56
Table 8: Standardized coefficients and classification results for discriminant models	59
Table 9: Significant correlations between subject variables and performance variables	70
Table 10: Attitudes toward videotex applications	75
Table 11: Multivariate analysis of effects in solvable and unsolvable problems	119
Table 12: Univariate analysis of Search Method effects in solvable and unsolvable problems	120
Table 13: Univariate analysis of Search Task effects in solvable and unsolvable problems	121
Table 14: Univariate analysis of Search Method x Task interactions in solvable and unsolvable problems	122
Table 15: Two-tailed tests of significance for multiple comparisons	123
Table 16: Multivariate analysis of effects in solvable problems	125
Table 17: Univariate analysis of Search Method effects in solvable problems	126
Table 18: Univariate analysis of Search Task effects in solvable problems	127

Table 19: Analysis of variance for regression of retrieval successes on spatial memory scores	128
Table 20: Analysis of variance for regression of pages accessed on ideational fluency scores	129
Table 21: Analysis of variance for regression of retrieval successes on spatial memory scores (Hierarchical search)	130
Table 22: Analysis of variance for regression of pages accessed on spatial memory scores (Hierarchical search)	131
Table 23: Analysis of variance for regression of pages accessed on ideational fluency scores (Keyword search)	132
Table 24: Multivariate analysis of differences in ratings between search methods	133
Table 25: Non-parametric tests of association between subject characteristics and pre-exposure attitudes toward videotex	134

LIST OF FIGURES

	Page
Figure 1: Outline of experimental plan	41
Figure 2: Mean number of retrieval successes as a function of search task	50
Figure 3: Mean number of pages accessed as a function of search task	52
Figure 4: Mean search times as a function of search task	54
Figure 5: Mean time per page as a function of search task	57
Figure 6: Six-group scatterplot with reference to Discriminant Functions 1 and 2	61
Figure 7: Summary comparison of search-method performance characteristics in unsolvable problems	63
Figure 8: Summary comparison of search-method performance characteristics in unipartite problems	65
Figure 9: Summary comparison of search-method performance characteristics in multipartite problems	66

INTRODUCTION: A RATIONALE FOR KEYWORD DIRECTORIES

The starting point for this paper is a discussion of the key problem areas which have been encountered in designing effective and efficient information-retrieval procedures for videotex. This is followed by a review of the attempts which have been made within the general constraints of keypad-only input to remedy existing procedures and to develop alternatives. A rationale is then presented for experimentation with keyword directories.

Gutenberg's New Clothes

As a public utility, the relatively inexpensive, mass medium of electronic information and transaction services known as interactive videotex¹ is still a novelty in Europe and comparatively untried in North America, and yet it is being hailed as an extraordinary innovation in communications services with the potential to revolutionize society.

Videotex is to the '80s what television was to the '30s and telephone to the 1870s. With the potential to change our leisure hours and even the way we do business, videotex is the communication system of tomorrow. (Wilson, 1980, p. 76)

The basics of our social life are going to be changed to a degree that they have not since the well-born German goldsmith, Gutenberg, began that mysterious ten-year process that led eventually to the creation of quickly reproducible,

1. For an explanation of "interactive videotex" and other related terms, see the glossary of videotex terms in Appendix A.

absolutely similar, metal type — letting any handy wine-press be turned into a bookmaking machine and putting more than a little panic into those with heavy investments in monasteries and scriptoria.

All information in all places at all times. The impossible ideal. But the marriage of computers with existing communications-links will take us far closer to that goal than we have ever been. (Godfrey, 1980, p. 1)

Children who grow up in a world in which they can access virtually any extant courseware from their living room, and in which a much greater proportion of learning is self-motivated, will end up with their cognitive processes organized differently from people of previous generations, whose learning environments have been far from optimal. They will be smarter, more knowledgeable, more skillful. They will be better equipped to solve problems in many areas of human endeavour. H. G. Wells thought that the fate of the human species would be determined by the outcome of a race between education and catastrophe. Developments in the past few years suggest that catastrophe is winning, but videotex-aided learning has the potential to tip the odds in favour of education. (Muter, Treurniet & Phillips, 1980, p. 323)

As extravagant as these claims are, they serve to highlight the perceived significance of videotex, as does also the current of optimism that runs like a thread through the published proceedings of recent videotex conferences (4th International Online Information Meeting, 1980; Proceedings, Inside Videotex, 1980; Transcript of Viewdata '80, First World Conference on Viewdata, Videotex, and Teletext, 1980; Videotex '81 International Conference and Exhibition, 1981). The truly telling point though is the intense competition among the leading nations in videotex technology — Britain with Prestel, France with Teletel, Canada with Telidon — to win the favours of the very large, and as yet unexploited, American market. This is especially evident in the implicit and explicit comparisons being made with respect to graphics protocols; but the great "graphics war" between Telidon, with

its alpha-geometrics, and the rest of the world, with its alpha-mosaics, has been declared by one commentator a non-issue:

Graphics, graphics, who has got the graphics? "We all do, we all do!" Remember the great tailfin wars in the automobile industry in the 1950s? In 2010, we will look back on the great videotex graphics wars of the 80s with the same bemusement! . . . Graphics and colour of themselves are of value only to the extent that they convey information or entertain. (Gaujard, 1980, p. 107)

At a more general level, Gaujard argues persuasively that it is not the technological wizardry of videotex that will attract paying customers but rather it is the variety and quality of services offered over the medium that will be the crucial determinants of public acceptance. A similar point of view, based on field-trial feedback, has been expressed elsewhere (QCLC, 1980). Consequently, not all observers have fallen under the spell of videotex enchantment, and several have been deliberately cautious in their reviews of Gutenberg's new clothes. The assumed mass-market appeal of videotex has, for instance, been appraised by INFOMART'S director of database publishing as follows:

Unfortunately, we have to acknowledge that the various videotex systems are still only interesting pioneer services. Most of the enthusiasm for selling information from videotex databases has sprung from our technological imagination and not from market analysis. (Mauerhoff, 1981, p. 414)

To redress this information gap, a number of market trials have been launched across Canada -- the Telidon Project in Ontario, Project Ida and the Elie Project in Manitoba, Project AGT/Telidon in Alberta, Project Mercury in New Brunswick, Project Vista and Telidon 2 in Quebec, and the Videotex Project in British Columbia (Feeley, 1981).

Let us assume then that the question of what to do with this hybrid technology will be answered fairly quickly. Let us even allow that videotex may eventually have the capability to deliver "all information in all places at all times" (Godfrey, 1980, p. 1). The residual questions, those bearing on the effectiveness and efficiency of videotex as an information utility, will still require answers. Specifically, how do videotex systems perform when put to an empirical test? What actually happens when a keypad is placed in the hands of an ordinary citizen and he/she is asked to operate the system? Are the contents and services of some remote but extremely powerful computer really at one's fingertips, "only the push of a button away," as has been claimed (CBC, 1982)?

The contents of videotex are not, in fact, as accessible as currently hoped, and it is the complexity of interaction between human and videotex machine which constitutes perhaps the most pressing problem area for videotex researchers. It is not too difficult to see why. If the advertised services of videotex remain relatively inaccessible, or awkward to use, videotex will not be perceived by the general public as a satisfactory utility and therefore will not sell. Hence a major focus of interest for videotex developers and researchers has been the optimization of interface characteristics for ordinary users. While some of the ergonomic questions of keyboard/keypad design have, to a large extent, been puzzled out by antecedent research on interactive computers, a great deal of territory remains to be explored in the interdependent areas of software characteristics, task characteristics,

and user characteristics (see, for example, Dillon & Tombaugh, 1982).

That the state of knowledge in these areas is still weak is made unequivocally clear by this skeptical review of available videotex systems:

The available systems have various degrees of flexibility ranging from the hierarchical, directed graph indexing of first generation Telidon, constrained by the current database design, through Prestel with its directed graphs which appear hierarchical but are less constrained, to the STAR system which provides for a rich variety of pre-coordinate indexes including keywords which point to single response pages. Such systems, particularly Telidon and Prestel, mimic existing paper based indexes and suffer from most of the same disadvantages with respect to multiple access-points, multiple user views and intolerance of variant (user) search strategies.

The present generation of videotex systems have [sic] a long way to go if they are to equal the propaganda of their creators These systems are pale imitations of an "electronic library". (Ball, 1981, p. 14)

The Essential Problem: Information Indexing and Retrieval

In many respects, the retrieval of information from computerized databases such as those developed for interactive videotex can be likened to the popular fantasy-adventure game "Dungeons and Dragons" -- albeit a much less colourful and elaborate version. Still, it requires only a modicum of imaginative licence to see in the information which is the target of a search an analog of the much sought-after treasure of D&D. Similarly, one can discern in the guise of the database indexer the features of a putatively benevolent Dungeon Master. The most striking parallel, however, and the key one, lies in the fact that, in both worlds, most of the action not only takes place in unfamiliar and uncharted territory but, as if one handicap were not enough, in the dark

as well. These metaphors are appropriate in the sense, firstly, that searchers rarely know beforehand the location of, or the route to, the information they seek, and, secondly, that whatever signposts may be encountered along the way in the form of indexing pages, they are not of equal illuminating power.

All of the major videotex systems — Prestel, CAPTAIN, Teletel, and Telidon -- have been built around hierarchical retrieval strategies (Dew, 1980; Inoue, 1980; Ball, 1981), essentially because this form of menu-driven dialogue is moderately flexible, is easy for inexperienced users to learn and operate, and is also the most economical of CPU time (Martin, 1973; Miller & Thomas, 1977; Phillips, 1980a; Shneiderman, 1980; OCLC, 1981; Thomas & Schabas, 1981). Typically, hierarchical searches involve the selection of increasingly-narrower classification terms from successive menus of terms until a selection is narrow enough to define one information document. Although the idea is simple enough, hierarchical systems have not proven to be simple or efficient in application. Despite the fact that users generally respond favourably to the concept of videotex, hierarchical retrieval tends to be perceived as slow (OCLC, 1981), tedious (Eissler, 1981), and cumbersome (Harashima & Kitamura, 1981). By objective performance standards, hierarchical searches are considered inefficient and error-prone, with most searches eating up roughly twice as many pages as optimally necessary and with approximately half the errors occurring in the top two levels of the hierarchy, so that searchers seem to end up in blind alleys almost as often as they find what they are looking for (Lee & Latremouille, 1980;

McEwen, 1980; Stewart, Cox & partners, 1980; Latremouille & Lee, 1981; Tombaugh & Dillon, 1981; Tombaugh & McEwen, 1982). These findings undoubtedly account for allusions in the literature to the "mineshaft" structure of videotex (Lane, 1980; Winsbury, 1981; Williams, 1981). When one considers, moreover, that these deficiencies have been noted in relatively small databases, it is not surprising that even gloomier outcomes are being predicted for databases which aspire to encyclopedic status (Phillips, 1980b; Jensen & Ball, 1981).

Improving Information Accessibility

Attempted solutions to the inherent problems of inflexibility and ambiguity in hierarchical retrieval generally fall into one of two categories: upgrading the operating features of hierarchical search methods, or developing alternative approaches. In the former category, we find such innovations as (1) the cross-referencing of category terms in several branches of the hierarchy and (2) the chaining together of logically related documents at the hierarchy's end-points to allow for the scanning of successive documents without recourse to an intervening menu page (Williams, 1981; Williamson, 1981). Furthermore, in an attempt to reduce the ambiguity of broad classification terms in the top levels of the hierarchy, brief descriptors have been appended to the terms, with some success demonstrated in reducing errors (blind alleys) by one third to one half (Latremouille & Lee, 1981). A somewhat different approach to getting the most out of a hierarchical system is being explored by ZOG researchers (Robertson, McCracken & Newell, 1981).

In what might be figuratively regarded as the muscular version of hierarchical systems, ZOG -- a large-network, rapid-response system -- generates extremely fast responses (0.10 seconds in one version, 0.25 in another) and relies on either a touch screen or single-character keystroke in order to accelerate the process of traversing index pages to a degree which will be satisfactory even to expert users. But while the tedium of hierarchical accessing has been relieved to some extent, Robertson and his colleagues readily acknowledge that other, more universal, human-factors problems have persisted. These include: disorientation in the database, failure to read the information in frames, and short-term memory limitations. Thus, while the ZOG system does offer several desirable performance features, they are not immediately transferable to commercial videotex, given the inability of existing telephone line to achieve the high signal-transmission rates required.

The fundamental weakness of hierarchical systems derives from the fact that any hierarchy constitutes only one of many possible ways of categorizing information, whereas any datum of information can in fact be included in as many logical categories as the mind can invent. Thus a category considered appropriate by the indexer for a given document will not always be the one explored by all users seeking that document. One approach to overcoming this inflexibility of hierarchical systems in the face of information's multi-dimensional nature signals a departure from purely hierarchical indexing even while ostensibly functioning as a supplement to it. This is the "alphabetical directory", which,

hypothetically, would be constructed as follows:

The source of index entries would be the database itself together with synonyms and near-synonyms added by the indexers. In theory, the entire database, including both index pages and document pages, could be indexed in such a directory. In practice, this approach presents problems and may hinder rather than aid effective access. Preliminary research and the examination of existing databases suggest that the content of subject directories should be limited to descriptors from the index pages, the main themes of the documents, and appropriate synonyms, near synonyms, and variant term forms. In other retrieval systems, it has been demonstrated that there is a degree of exhaustivity of indexing beyond which there is improvement in retrieval and beyond which additional indexing may adversely affect the efficiency of the system. (Williamson, 1981, p. 11)

What emerges from this solution is a system which anticipates keyword retrieval, but which, in view of the relatively non-specific level of indexing, actually represents only a slight loosening of the hierarchy's joints, comparable to what Meadow (1973) defines as "subject-heading indexing". In practice, therefore, alphabetical directories of this sort rarely lead directly to documents at the lowest levels of the hierarchy but, rather, to some intermediate level, at which point selection by menus of category terms is again resumed. As Hardy (1978) and Williamson (1981) point out, if actual documents and pages are used to generate directory terms, the directory runs the risk of becoming too big and unwieldy to work efficiently. Thus the directory terms developed for the Prestel system and an experimental Telidon system, were relatively simple -- category names, names of information providers, and their variants (Rimmer, 1979; Stewart, Cox & partners, 1980; Williamson, 1981; Tombaugh & McEwen, 1982).

How do alphabetical directories actually perform? In a comparison of hierarchical and alphabetical-directory indexing on Prestel, no major differences were detected between methods in search times, total pages viewed, or blind alleys (Stewart, Cox & partners, 1980). Both methods performed equally badly. Likewise, in a partial replication of this study on Telidon, the alphabetical directory failed to demonstrate any performance edge over the hierarchical index (Tombaugh & McEwen, 1982). One reason for this unexpected outcome may be that terms tended merely to reflect the underlying hierarchical structure as opposed to cutting across it with new and diverse relational links. In other words, the very simplicity and generality of the directory's terms may have been its undoing.

Keyword Indexing: An Alternative Approach

In light of some of the above-noted difficulties in information retrieval, at least one group of experts has concluded that "the cumbersome tree structure searching method which these videotex systems employ needs to be replaced with a more direct and efficient method of interaction" (Information Services Co-ordinating Group, 1981, p. 9). Schabas, Williamson and Eines (1981) are quite certain of the direction to be pursued -- keyword systems. Unlike hierarchical systems, which force the searcher to adapt his/her thinking to one dominant and palpable structure of relationships in the database, keyword systems allow for several links among database elements. Consequently, items can theoretically be retrieved by any of a number of routes. Also, less

time and effort are required of the searcher, since entry to the database is accomplished with a fairly specific term in contrast to the broad term mandatory for entry via a hierarchy. In an independent statement, however, Williamson (1981) was more cautious in advocating keyword indexing as the most appropriate substitute, her reservation seemingly stemming from the unforeseeable impact on ordinary users of implementing new, suitable software and hardware. It should be noted that, in this context, suitable hardware invariably means an alphanumeric keyboard; but among the more interesting videotex experiments have been those in which some attempt has been made to implement a keyword system while retaining the numerical keypad's assumed advantages as a user-familiar and low-cost input device.

Pollard and James (1980), for example, evaluated several schemes for numerically coding the letters of the alphabet so that simple terms from an alphabetical directory could be entered on a numeric keypad. In one format, each letter from A to Z was assigned a corresponding code from 01 to 26 which remained unchanged throughout the index.

The initial letter page consisted of all 26 letters with corresponding codes 01-26. Thus to access HOTELS, for example, after inputting 08 for H, a second level index page displayed all grammatically possible letters which can follow H, each with their own unique code. Thus 15 would be input for O and so on until the HOTELS page was found.

1st Page.

01:A, 02:B, 08:H, 26:Z
Key 08 to obtain 2nd Page.

2nd Page.

01:HA, 05:HE, 09:HI, 15:HO, 21:HU, 25:HY
Key 15 and so on to obtain HOTELS.

(pp. 11-12)

While this study succeeded in demonstrating that a 26-fold breakdown of the alphabet was faster and less error-prone than a 10-fold breakdown (the latter reflecting Prestel's conventional 10-choice limitation per page), the researchers conceded that alpha input by such a coding scheme could be annoyingly laborious. As it is, the term HOTELS requires six index pages just to get to the right category. To access information on topics such as RESTAURANTS or AIRLINE RESERVATIONS would quickly inflate the number of indexing pages required to unacceptable levels.

Taking a different tack, Hardy (1978) theorized that keypad numbers superimposed with groups of alphabetical characters, in the fashion of older telephone dial codes, might allow for the direct numeric coding of keywords. Several coding systems were developed with the joint objectives of avoiding the generation of more than one authorized keyword from any individual keying sequence while simultaneously keeping the number of keystrokes to a minimum. What ultimately emerged from testing as the optimum system (1.24 words generated per keystroke sequence) consisted of the following alphanumeric pairings --

A	BC	DE	FGH	IJKL	MN	OPQ	RS	TUVWXYZ
1	2	3	4	5	6	7	8	9

and an abbreviation-keying strategy which used as many initial letters of the entry terms as possible, namely: if the entry term was one word -- the first four letters; if the entry term was two words -- the first two letters of each; if the entry term was three words -- the first two letters of the first word and the initial letters of the other two; and if the term was four or more words -- the initial letters of the first

four significant words. Such an approach might first appear, from a human-factors perspective, to be error-prone and unworkable. Yet at least one information provider had implemented a reasonable facsimile of this indexing system on Prestel by the time Stewart and his colleagues (1980) evaluated the various accessing procedures available, and surprisingly, the limited keyword system reduced search times, total pages accessed, and blind alleys by 40% to 60% when compared against hierarchical and alphabetical-directory methods. It was not clear, however, whether keyword indexing per se or some other feature of the system had actually contributed to the success of the search method. One drawback to this approach, furthermore, is that without a printed directory of authorized keywords the searcher has no way of knowing what numerical codes to use (Hardy, 1978).

A Logical Progression: Keyword Directories

A retrieval procedure which obliges the user to refer to a printed directory undermines somewhat the purpose of an electronic information system. If system independence is desired, an on-line directory becomes an essential goal. But once an on-line directory of keywords is in place, the procedurally awkward method of accessing documents by numeric coding of keywords can be relinquished in favour of direct accessing by page number. For example, in the course of mulling over the possibility of abandoning the tree structure altogether, Thomas and Schabas (1981) suggested:

An intermediate approach would be to retire the tree structure into the background, to be brought out and

- displayed as an ordered, classified index when needed to assist the user. If pages referenced in the index were given titles, and each page retained its visible number, then the user's mental image might be that of a book: a book which he could open directly at a well known page, enter through [a] back-of-the-book word index, or through its elaborate table of contents that topically orders chapter and section headings. (p. 7)

This was the approach adopted in the keyword system developed for evaluation in this research. Database documents, which, for the most part, already possessed titles, were assigned an average of five keywords or keyword phrases, including general, specific, and variant terms. These were subsequently compiled in one alphabetical list, each keyword paired with the page numbers of the documents to which it pointed. For each keyword, an indexing page was created which consisted of the titles of all database documents to which each keyword referred and the corresponding, directly accessible page numbers. Calling up these document-title pages could then be accomplished by the method evolved for the alphabetical-directory index, except that initial entry to the alphabetical index would be through a 26-fold breakdown of the alphabet (Pollard & James, 1980).

Although the opportunity existed, with the use of computer terminals, to take advantage of full-keyboard facilities for keyword input, a 15-key keypad was selected as the most appropriate input device since most interactive videotex systems are based on keypad input (Mahoney, Demartino & Stengel, 1980; OCLC, 1980; Stewart, Cox & partners, 1980; Tydeman & Zwimpfer, 1981). It should be noted, incidentally, that, while practically any available terminal keyboard

can serve as a videotex input device, most do not support Telidon graphics and are not useable, therefore, in a Telidon network. More importantly, there has been no justification to date for acquiring a keyboard facility; the range of services currently being offered on videotex systems (Aysan, 1980; Fraser, 1980; OCLC, 1980; Eissler, 1981; Harashima & Kitamura, 1981; Tydeman & Zwimpfer, 1981; Williams, 1981) is simply not that sophisticated. From a marketing perspective, moreover, there are a number of pragmatic considerations which make the keypad more likely to be accepted by large numbers of inexperienced computer users. Not only is a keypad a more familiar instrument — comparable to push-button phones, digital television tuners, and the ubiquitous hand calculator — but it will presumably be easier to use and will cost less, factors which cannot be underestimated even by the gadget conscious: "Videotex was conceived with the general non-specialist market in mind. Ultimately this is the mass residential market, where low cost and general utility are at a premium" (Woolfe, Cox & partners, 1980, p. 57).

Notwithstanding the theoretical objections raised with respect to the indexing of individual information documents (Hardy, 1978; Williamson, 1981), it remains to be seen whether such directories do in fact become unwieldy and inefficient. Furthermore, the empirical limits of indexing systems supported by a numerical keypad have not yet been established, and keyword research of this sort might serve to indicate whether keypads are a good long-term bet as input devices. Thus, beyond the basic objective of determining the feasibility of keyword indexing

by means of an on-line directory, the major purpose of this study was to evaluate keyword retrieval against hierarchical retrieval within the general constraints imposed by numeric-only input.

This comparative evaluation of two search methods was expected to afford some insight into the interface characteristics of the particular human-machine combination called videotex. It was also expected that user performance in this system, as in other interactive systems, would be mediated by other variables besides search method. For example, the major classes of variables identified by Martin (1973), Mason and Mitroff (1973), Miller and Thomas (1977), Ramsey and Atwood (1980), Shneiderman (1980), and Benbasat, Dexter and Masulius (1981) include not only the various hardware and software combinations which constitute the interface characteristics but also focus on the characteristics of the interactive task and the characteristics of the user. Hence it was considered essential that a just comparison of retrieval methods also address at least some elements of these variable domains.

Search-Task Characteristics

With respect to retrieval-task characteristics in videotex systems, at least one investigation of user performance in uncertain-information situations has been concluded under a hierarchical search condition (Whalen & Latremouille, 1981). No one has yet compared this kind of retrieval task across search methods, nor has anyone explored search performance in situations where more than one document is required to

solve the retrieval task. On the basis of the more numerous relational links attainable in a keyword system relative to a hierarchical system, it could be anticipated that a keyword system would provide better meta-information for both multiple-solution and unsolvable tasks. Even in tasks involving the retrieval of single documents, a keyword system could again be expected to out-perform a hierarchical system, given that keywords grant database entry with more concrete and specific terms and open up several routes to target information.

The second major objective of this research was, therefore, to compare hierarchical indexing and keyword indexing across different types of retrieval tasks. Accordingly, the experimental test-battery consisted of three classes of retrieval problems: unipartite solution, multipartite solution, and unsolvable. Unipartite-solution problems were characterized by the fact that only one information document and page in the database could provide the correct answer. Multipartite-solution problems required several (two to six) independently retrievable documents and pages to constitute a complete answer. Unsolvable problems posed questions which were plausibly answerable but could not, in fact, be solved by any existing document in the database.

Measuring Search Performance

The measures which have been used most commonly to evaluate and compare interactive information-retrieval systems are summarized in Table 1, categorized according to whether they are more inclined to tap the information-extraction power of the retrieval system or the effort that must be expended by the user. Conspicuous by its absence from the list of effectiveness indices is a direct measure of recall, "the ability of the retrieval system to uncover relevant documents" (Lancaster, 1968, p. 55). While search errors could be claimed to measure effectiveness indirectly, a direct measure seems much more appropriate. As Lancaster (1968) has pointed out, "Recall is obviously the most important requirement of the user of a retrieval system since his sole purpose in approaching the system is to obtain one or more documents useful in relation to his information need" (Lancaster, p. 55). This view has been echoed by Schabas, Williamson, and Eines (1981). Several researchers have provided general information about recall: Frankhuizen and Vriens (1980) report percentages ranging from 20% to 90% for correctly solved search tasks; Stewart, Cox, and partners (1980) cite a 79% success rate in Prestel research; Tombaugh and McEwen (1982) report 98.7%; Whalen and Latremouille (1981), 86.5%. But, in spite of its central importance, recall has not figured as a dependent variable.

While it may be safe to assume that recall is not an essential measure when the rate of success in retrieval is consistently high or

TABLE 1

Performance measures used in previous research

Measure	Researchers
<u>Effectiveness</u>	
Search errors	Frankhuizen & Vrins (1980) Lee & Latremouille (1980) McEwen (1981) Stewart, Cox, & partners (1980) Tombaugh & McEwen (1982)
Data-entry errors	Frankhuizen & Vrins (1980) Pollard & James (1980)
<u>Effort</u>	
Pages accessed	Frankhuizen & Vrins (1980) Robertson, McCracken, & Newell (1981) Stewart, Cox, & partners (1980) Tombaugh & McEwen (1982) Whalen & Latremouille (1981)
Search time	McEwen (1981) Pollard & James (1980) Robertson, McCracken, & Newell (1981) Stewart, Cox, & partners (1980) Tombaugh & McEwen (1982)
Time per page	Robertson, McCracken, & Newell (1981) Stewart, Cox, & partners (1980)
Keystroke frequencies	Robertson, McCracken, & Newell (1981) Tombaugh & McEwen (1982)

practically invariable, such an assumption becomes untenable if retrieval success is markedly imperfect or highly variable, as it could be under natural conditions. Focussing exclusively on indices of search effort, or work, could easily lead to erroneous conclusions if, for example, the fastest search method is also the least productive in terms of extracting relevant documents from the database. Similarly, it should be apparent that emphasizing recall at the expense of work measures would be an equally grave mistake, especially when one considers that it is always possible to achieve 100% recall by retrieving every document in the database (Lancaster, 1968). Obviously, some balance must be struck between measuring retrieval effectiveness and retrieval effort.

The indices conventionally adopted to meet this requirement have been recall ratio, the ratio of relevant documents retrieved to known relevant documents in the database, and precision ratio, the ratio of relevant documents retrieved to the total number of documents retrieved (Doyle, 1975; King & Bryant, 1971; Lancaster, 1968; Meadow, 1973; van Rijsbergen, 1975). Applied to industrial-scale information services, these indices are calculated exclusively from the products of the search. They do not take into account the intermediate processing, which is invisible to the inquirer. In interactive systems though, the query process itself springs into high relief. Hence the page, as the means and object of the search, becomes the more suitable and calculable unit of measurement. Still, there is a major obstacle to the wholesale adoption of recall ratio and precision ratio as measures appropriate to

interactive systems. Unsolvable problems cannot be meaningfully accommodated by either one since the absence of relevant documents will always result in constant values. An uncomplicated expedient would be to dispense with ratios and retain the simple count of relevant pages retrieved as the index of recall and the totals of pages accessed and search times as the indices of work performed. Table 1 indicates that these latter two, at least, are in common use.

For this study, the count of retrieval successes, or information hits achieved in a search task, was chosen as the recall measure, and the number of pages accessed and search time were selected as precision measures. Although search time was expected to correlate positively with the number of pages accessed and consequently provide some redundant information, some portion of the variability in search time might conceivably be attributed to the different cognitive demands made by each search method, and therefore would be a source of unique information as well. Time per page was also included in the inventory of measures in recognition of the possibility that information-processing demands might vary across search methods. Keystroke frequencies were considered unlikely to show differentiated performance since both search methods being tested were menu-driven and were to be operated by the same numeric-input conventions. The measurement of user error, whether cognitive or data-entry, was a more problematic issue. While search errors of the cognitive kind could be readily operationalized as "blind alleys" taken in the search/decision tree, the calculation of such errors from raw data was foreseen as being

extremely complex in multipartite tasks and meaningless in unsolvable search tasks. Additionally, no certain means could be devised for separating out data-entry errors undetected by the user. However, the principal logic for discarding the plan to measure error was that the consequences of both cognitive and mechanical errors would be manifested in larger totals for pages accessed and longer search times. In other words, error information would be indirectly available in another form.

In summary, the foregoing considerations led to the selection of four measures of search performance: one measure of recall -- the count of target pages successfully retrieved; and three measures of work -- the count of pages accessed, search time, and time per page.

User Characteristics

Among the user characteristics which have been studied in automated information systems, cognitive style, or cognitive ability, has been singled out as especially relevant to interactive systems (Bariff & Lusk, 1977; Benbasat, Dexter & Masulius, 1981; Doktor, 1976; Mason & Mitroff, 1973; Shackel, 1980). Unfortunately, only very tentative conclusions can be gleaned from the literature on cognitive abilities in interactive systems. In two studies (Bariff & Lusk, 1977; Benbasat, Dexter & Masulius, 1981), Witkin's Group Embedded Figures Test showed both non-differentiating and differentiating effects in the performance of Management Information System users. In the latter study, high scorers, characterized as high-analytic subjects, out-performed low

scorers on a number of information-gathering and decision-making tasks. More recently, a series of Bell Laboratories studies has examined the relationship between learner characteristics and performance in text-editing tasks (Egan, Bowers & Gomez, 1982; Egan & Gomez, 1982; Gomez & Egan, 1982). Spatial memory, as measured by the Educational Testing Service's Building Memory Test, was found to correlate systematically with measures of task-completion time and error frequencies. To the surprise of the researchers, verbal skill correlated only weakly with one of the dependent variables. Naturally, one wonders if there might be a link between the differentiating abilities identified in these seemingly disparate studies and if such a connection might have implications for videotex systems.

If there is a common element, it would seem, from the available evidence, to be a spatial factor. Witkin and his associates (1974, 1979) have limited the generality of a disembedding ability in perception, as measured by the Embedded Figures Test, to performance on spatial-visualization tasks. The Building Memory Test is also heavily loaded with a spatial factor. In the research in which these tests were used, test scores were found to correlate positively with qualitatively different dimensions of performance in computer systems. This suggests that spatial ability might mediate performance in interactive computer systems in some general way. Theoretical support for this inference is offered by Doktor's speculations (1976) on the possible effects of lateralization of cerebral function on cognitive performance in computer systems. He has proposed that performance may be mediated by the user's

ability to develop and retain an appropriate mental model for the system in use. In some situations, the appropriate model may be a verbal/analytic one; in others, a spatial/relational one. If a strong spatial effect were to be detected in videotex systems, it would provide confirmatory evidence that retrieval tasks are primarily navigational in nature and that they are best solved in what Jaynes calls the mind's "analog space" (1976).

Whether spatial ability affects performance in videotex systems is not known. In their conjectural arguments for the provision of navigational aids in hierarchical systems, Lochovsky and Tschritzis (1981) seem to take it for granted that the spatial demands of such an indexing system are considerable. It has also been noted that disorientation, an effect normally associated with spatial problems, was a recurring phenomenon in the ZOG system (Robertson, McCracken & Newell, 1981), which is a hierarchical information-retrieval system similar in many respects to commercial videotex. Weak spatial ability could have been partially responsible for this effect.

In order to establish whether spatial ability does mediate performance in videotex systems, a subsidiary objective of this research was to calculate the degree of association between spatial ability, as measured by the Building Memory Test (Ekstrom, French, Harman & Dermen, 1976); and the various performance measures obtained under both search procedures. The psychological meaning of the cognitive factor tapped by this scale has been defined by Ekstrom et al as "the ability to remember

the location of figural material" (p. 109), and has been identified, in Guilford's well known structure-of-intellect model (Guilford, 1967), with memory for figural units and relations.

Despite the prospect of a demonstrable relationship between spatial ability and videotex performance, a reasonable alternative cannot be overlooked. There is also a possibility that verbal ability may be a factor in information retrieval. The content of videotex is, after all, primarily textual, and keyword searches do not appear to be intrinsically navigational tasks. In fact, when these search procedures are viewed from a problem-solving perspective (Ramsey & Atwood, 1980), the generation of relationships between knowns and unknowns appears to be the central process in keyword searches, whereas the recognition of such relationships seems to predominate in hierarchical searches. From a functional point of view, some hierarchical features will be present in the keyword system -- since it is menu-driven -- but verbal abilities bearing specifically on the production of appropriate entry terms seem more likely to dominate. This kind of performance could be considered as occupying the divergent-production-of-semantic-units cell in Guilford's model (Guilford, 1967), so that a test tapping this performance domain, such as the Ideational Fluency scale of the Comprehensive Ability Battery (Hakstian & Cattell, 1976), might differentiate levels of performance in the keyword search procedure. The psychological meaning of the factor measured by the Ideational Fluency scale is given by Hakstian and Cattell (1976) as follows:

This ability is concerned with producing ideas about a given topic rapidly and without much attention to quality
The FI test of the CAB is of the "attribute-listing" type,

in which examinees must list as many adjectives as they can, in a fixed time, that could be applied to a given thing. This ability to retrieve learned ideational material when it is needed is important in many situations, and, as with Spontaneous Flexibility (Fs), appears to be a facet of originality. (p. 7)

As in the case of spatial ability then, a parallel research aim was to determine if there is a relationship between verbal ability and performance in videotex systems. For the reasons mentioned, it was anticipated that a verbal factor would be of greater consequence in keyword retrieval.

User Evaluations

Because public acceptance of interactive systems is premised on favourable subjective reactions to such a service, irrespective of objective performance criteria (Ball, 1981; Cuff, 1980; Feeley, 1981; Gaujard, 1980; Martin, 1973; Moran, 1981; Nickerson, 1981; OCLC, 1981; Shneiderman, 1980; Williamson, 1981), two kinds of user response seemed necessary to complement performance comparisons. The first pertains to whether exposure to a videotex system can affect attitudes toward the medium in a positive way. There is some evidence from market trials that it does (Eissler, 1981; Harashima & Kitamura, 1981; OCLC, 1981). To confirm this effect, a scale was devised to measure user attitudes to a range of hypothetical videotex services before and after experimental exposure.

The second essential aspect of user reaction concerns the point at

which performance differences between search methods become psychologically meaningful to the user. From a marketing point of view, the detection of significant performance differences will not be very useful if these differences are not translated into a rating spread. In the same vein, the extent to which performance differences predict search-procedure preferences is critical information to those who must make decisions about system development. To measure user impressions of search-procedure performance and utility, a rating scale was constructed on theoretical grounds from a review of the literature on interactive systems (Martin, 1973; Mason & Mitroff, 1973; Miller & Thomas, 1977; Cuff, 1979; Card & Moran, 1980; Ramsey & Atwood, 1980; Shackel, 1980; Shneiderman, 1980; Moran, 1981; Nickerson, 1981; Williges & Williges, 1982). Constructs identified in these references and deemed relevant to user satisfaction were reduced to a short list by clustering terms according to commonality of meaning and subsequently assigning a single descriptive term or phrase to each theoretical construct. Rating levels were then provided for each sub-scale.

Concerning user evaluations, it was expected that as a result of exposure to a videotex system a more favourable attitude would develop among users. Secondly, it was expected that the search procedure which performed better by objective standards would also be rated more highly.

Summary of Research Hypotheses

The following hypotheses were tested in this research:

- (1) concerning the effectiveness of a menu-driven hierarchical versus a menu-driven keyword search procedure, that the keyword system would out-perform the hierarchical system, retrieving more relevant information in less time and with less effort;
- (2) concerning the effects of task variety on search performance, that keyword searches would require fewer pages and less time than hierarchical searches (i) to retrieve more target information in solvable unipartite and multipartite tasks and (ii) to confirm the non-existence of target information in unsolvable problems;
- (3) concerning possible relationships between individual abilities and the differential cognitive demands of each search procedure, that (i) a relationship would be demonstrated between spatial ability and search performance under the hierarchical, or both, search conditions, and (ii) a relationship would be demonstrated between verbal ability and keyword retrieval;
- (4) concerning comparative degrees of user satisfaction with the search procedures tested, that the keyword procedure would be rated as superior; and
- (5) concerning the utility of videotex as an information medium, that exposure would increase the favourability of user opinions.

METHOD

The Experimental Information System

An interactive, computerized information system, based on contents of the 1982-83 edition of the Saint Mary's University Academic Calendar, was used to simulate a videotex information service. The database developed for this study contained approximately 1225 videotex pages, of which about 700 were information pages. Approximately 25 pages were instructional or user-aid pages while the remaining 500 or so were indexing pages for the hierarchical and keyword search systems.

Page formatting, operating characteristics, and special features of the experimental system mirrored commercial videotex systems. Page numbers were coded for display in the top right corner of each page while the bottom two to four lines were reserved for keying directions, user input, and error messages. Executable commands consisted of either a number string (representing a menu selection or page number), or a function, and the ENTER keystroke. Following Dillon's and Tombaugh's recommendation (1980), one of the keys was defined and labelled as a HELP function. This key called up a one-page precis of system-operating conventions and error-correction procedures, while subsequent use of the BACK UP function returned the user to the page from which HELP was requested. Other features included direct accessing of information pages by page number (four digits) and, at the lower levels of the hierarchy, a browsing analog in the form of sequential accessing of

related documents.

Errors could be corrected in several ways. If a numerical or function entry error was detected before the ENTER key was pressed, the incorrect entry could be erased with the DELETE key. If an error was not detected until after the ENTER key had been pressed, or if the user simply wished to retrace his/her search route, the BACK UP key could be used to return to a previously displayed page and a new selection keyed. The capacity of the stack stored for retracing purposes was 25 pages; Prestel allows three back-up steps (Dew, 1980), Telidon ten (Ball, 1981). Errors due to illegal commands (e.g. unauthorized menu selection or illegal page number) resulted in the message "SORRY, I DON'T UNDERSTAND, PLEASE TRY AGAIN." If an attempt was made to access directly a page which did not in fact exist in the database, the message "SORRY, I CAN'T FIND PAGE XXXX" was displayed. Finally, a RESTART command (0 ENTER) allowed the user to return to the root page of the information system at any time and from any page.

Hierarchical search procedure

Hierarchical searches, or tree searches as they were identified for experimental subjects, began with a menu display of the six broad information categories according to which the database had been organized. (See Appendix B for an illustrative hierarchical search.) Brief descriptors were appended to the broad category terms in order to reduce the possibility of conceptual errors at upper levels of the

hierarchy (Stewart, Cox & partners, 1980; Latremouille & Lee, 1981). At this stage, the searcher had to decide which category, or branch of the information tree, was most likely to lead to the answer to his/her query and then had to enter the number of the menu item which corresponded to his/her choice. A new menu was then displayed and the selection process continued until, if no cognitive errors or undetected keystroke errors were made, the desired document appeared on the screen. Thereafter, the searcher was referred to a new-topic page which permitted access to one of several higher-level or lateral menu pages, or exiting from the system.

Keyword search procedure

The first page displayed upon initiating a keyword search was a 26-fold breakdown of the alphabet. (See Appendix C for an example.) The user then selected the alphabetic range which was believed to contain the search term. This resulted in an alphabetical listing of all indexing terms in the chosen range which had been authorized for the database; if more than one page was required to list the terms in the selected range, information to that effect and keying instructions were shown at the bottom of the page. If the desired term or a variant was not listed, the searcher had to back up and try a synonym or some other entry term. Otherwise, selection of an authorized keyword resulted in a display of the titles of all relevant documents, the title of each document paired with its directly accessible page number. Although the keyword system was fully integrated with the hierarchical structure,

thereby saving the space which would otherwise be required for a parallel database, entry to the hierarchy occurred only at the bottom so that the search procedures were functionally independent. Browsing and exiting, however, were effected in the same way as in the hierarchical system.

Apparatus and Instrumentation

The computer used to host the calendar information system and to record performance measures on-line was an LSI 11/23 processor with 256 Kilobytes of memory and with 8.8 Megabytes of storage available on a fixed, high-speed Winchester disk. Three Visual 50 terminals, supported by the RT-11 and TSX-Plus time-sharing operating system, were available for individual or concurrent subject experimentation. Certain keyboard modifications were made to emulate keypad-only videotex systems. These included:

- (1) covering the keyboard proper, the DELETE key excepted, with a fibreboard shield so that only the keypad and DELETE key were available for subject use; the cover also displayed a reminder that the experimental convention for indicating retrieval successes was "HIT ENTER" while that for terminating a search was "98 ENTER".
- (2) redefining and labelling certain non-numerical keys on the keypad with the specific functions which they performed, i.e.

HELP -- to access the HELP page, HIT -- to record a retrieval success.

The essential work of accessing the database, displaying pages at the terminal, monitoring user input, and collecting performance data was accomplished by a FORTRAN program. Each page, whether an indexing page or an information page, was stored in a unique file on the Winchester disk, thus allowing for rapid access and individual page modification and editing when required. Typically, each file contained the information to be displayed for the user and the list of authorized responses for that page. For each authorized response there was also a corresponding page number to link the user's entry to the next page (file) to be retrieved.

Performance measures were collected on-line in a data file which consisted of a log of all transactions between each subject and the information system. Specific items included were: a list of all system pages accessed; a record of the number of error messages displayed on any one page; a record of the elapsed time on each indexing or information page; a record of the elapsed system time between pages; a record of all data-entries, including information HITS, made on any page; and a record of the corresponding problem number for each data line.

The questionnaire shown in Appendix D was used to identify subject characteristics of interest, i.e. age, gender, handedness, academic

major, and level of experience with interactive systems. Attitudes toward videotex, both pretest and post-test, were measured by a four-item scale (see Appendix E) which permitted the subject to categorize his/her reaction to each of a variety of hypothetical videotex applications as (1) Unfavourable, (2) Unsure or no opinion, or (3) Favourable. The scale used to measure and differentiate subjects according to verbal ability was the Ideational Fluency (FI) test from the Comprehensive Ability Battery published by the Institute for Personality and Ability Testing. The scale used to measure subject differences in spatial memory was Educational Testing Service's test of Building Memory (MV-2). The psychometric characteristics of these tests are shown in Table 2. Subjective ratings of the search procedure used were measured on ten, 7-point, Likert-type scales (see Appendix F).

Pilot Study

Pilot research with a group of 12 student volunteers was conducted for the purposes of debugging experimental procedures and identifying any major structural flaws (e.g. category ambiguity) or omissions (e.g. obvious keywords) in the information system. With regard to search strategies adopted by subjects in the two search conditions, informal protocol analysis revealed that, in the hierarchical condition, subjects left to their own devices usually selected terms that were optimally compatible with the hierarchical index, that is, broad terms; but, in the keyword condition, subjects tended to generate search terms that were far from optimal with respect to the highly specific keyword index,

TABLE 2

Reliability and validity estimates
for tests of spatial memory and ideational fluency

Test	Split-half reliability	Factorial validity
Building Memory	.80	—
Ideational Fluency	.78 - .84	.88

that is, they too selected broad terms. Consequently, experimental instructions and example problems were modified to make explicit the most appropriate strategy for the search method being used: in the case of a hierarchical search, the selection of a broad search term; and, in the case of a keyword search, the generation of a narrow search term. Subsequent inspection of the pilot data suggested that this intervention may have been a factor in the reduction of search times and pages accessed in the keyword condition by as much as 50%.

Search tasks in the pilot study were a mixture of unipartite and multipartite solvable problems, but several subjects stated that they were unsure as to whether they were supposed to look for one page of information or more than one page. To correct this ambiguity, the phrasing of search tasks was standardized as follows: unipartite tasks all began with the phrase "Find the page", whereas multipartite tasks began with the phrase "Find as many pages as you can." With regard to the content of search tasks, care was taken to avoid excluding information which might be disadvantageous to one search method simply because of gaps in the general knowledge of the subject. For example, rather than stating a task in the form, "Find as many pages as you can that describe courses on Immanuel Kant", the problem was written "Find as many pages as you can that describe courses on the philosophy of Immanuel Kant". While the keyword searcher would not benefit from the extra information if "Kant" were the search term, the hierarchical searcher would not know where to begin if he/she did not know that Kant was a philosopher.

The prototype rating instrument was a 5-point scale on which subjects were required to indicate their degree of satisfaction with each performance dimension. Pilot ratings reflected negligible variability. In both conditions, performance dimensions were rated either as satisfactory or very satisfactory; no differentiation was evident between conditions. While the most appropriate solution in endeavouring to obtain a more discriminating measure of user satisfaction would have been to make Search Method a within-subjects factor in the research design, this option was impracticable. As a compromise, an attempt was made to improve the sensitivity of the rating instrument by increasing the rating levels from five to seven and by anchoring the levels with behavioural descriptors.

The final item of interest to emerge from the pilot study was that search time was highly correlated with the number of pages accessed, $r = +.90$, $p < .01$. Although not totally unexpected, this result did suggest that one measure was highly redundant.

Experimental Design

The model selected for this research was a two-factor mixed design with Search Method serving as the between-subjects factor and Search Task as the within-subjects factor. There were two levels of search procedure (hierarchical and keyword) and three levels of search task (unipartite-solution, multipartite-solution, unsolvable).

Subjects

With academic credit offered as an incentive for participation, 42 subjects (27 male, 15 female) were recruited from the Introductory Psychology classes at Saint Mary's University. The academic majors of the participating students represented a cross-section of the University's Faculties and programs. Ages in the sample ranged from 17 to 29 with a mean of about 20. About half of the students had significant prior computer experience, that is, at least one course in either programming or some other computer application. Only two students had been exposed to videotex. All students had previous experience in using the print version of the University's academic calendar. The distribution of subject characteristics on these and other variables of interest is summarized in Table 3.

Subjects were randomly assigned to either a hierarchical or keyword search condition but attempted all three kinds of search task.

TABLE 8

Subject-group composition
according to pre-measures

Group	Age	Computer experience		Gender		Handed		Spatial ^a score	Fluency ^b score
		Y	N	M	F	L	R		
Hierarchy ^c	19.4	11	10	13	8	3	18	15.3	48.6
Keyword ^c	20.3	9	12	14	7	3	18	17.5	48.6
Combined	19.9	20	22	27	15	6	36	16.4	48.6

^a Raw scores out of a maximum of 24.

^b Percentile scores.

^c n = 21.

Procedure

Figure 1 schematically illustrates the major steps in the experimental plan.

Pretest

Prior to the experimental session, the general nature of the experiment and the phases involved were explained to subjects. They subsequently completed a background-information sheet, pretest attitude questionnaire, and tests of spatial memory and ideational fluency.

Experimental session

Instructions concerning the operation of the information system and the requirements of the retrieval tasks were read to subjects (see Appendix G). Each subject was (1) familiarized with the keypad, (2) guided through a practice retrieval problem designed to illustrate the use of all function keys, and (3) given a practice retrieval problem. As recommended by Cuff (1979), the underlying intent of instruction was to impart a conceptual model of the search strategy appropriate to the assigned retrieval method, and to give each subject incremental hands-on experience with the procedural aspects of information retrieval.

Subjects attempted a battery of 20 information-retrieval problems, consisting of eight solvable unipartite tasks, eight solvable

FIGURE 1

Outline of experimental plan

(1)

 Collect background data and
pre-test videotex attitudes

(2)

 Measure spatial memory
and ideational fluency

(3)

 Randomly assign subjects
to search method

HIERARCHICAL GROUP

(4)

 Instruct subjects
in system operation
and search procedure

(5)

 Administer battery
of search tasks

(6)

 Measure performance

KEYWORD GROUP

(4)

 Instruct subjects
in system operation
and search procedure

(5)

 Administer battery
of search tasks

(6)

 Measure performance

(7)

 Post-test videotex attitudes

(8)

 Obtain ratings
of search method used

multipartite tasks, and four unsolvable unipartite tasks; the order of problem presentation was randomized for each subject. For each problem, subjects were required to locate and display the information document and page in the database which contained the answer to the question posed. In the case of problems for which several documents were necessary to supply a complete solution, subjects were required to locate and display as many as they were willing and able to find. Once having found a target page, subjects were required to record a "HIT" on that page with the key designated for that purpose. Search tasks were structurally independent of each other in as much as each search began on the root page of the information system and no task could be re-attempted once a new search was under way. Problems were presented, one at a time, on index cards. This arrangement was deemed preferable to on-line problem presentation, which raised the possibility of confounding performance variability with short-term memory effects.

The collection of performance data began as soon as the subject entered the first problem number with the appropriate keystrokes. For the purposes of analysis, however, relevant data were deemed to begin at the point in each problem at which the first considered choice had to be made. In the hierarchical condition, this was on the page which listed the general information categories. In the keyword condition, this was on the page displaying the 26-fold breakdown of the alphabet. The reason for this procedure was that some subjects in the pilot study automatically keyed in the start of a search before reading the task statement, whereas others first read the task statement, deliberated for

a few moments, and then keyed in the start-up sequence for a search. Eliminating these preliminary elements from the calculations of search time helped standardize this measure.

When the subject believed that the only or last target information page for a particular problem had been found, he/she was required to key "98 ENTER". This command terminated the current search task, prompted the subject to enter the number of the next search task, and cued the computer to display the root page of the information system. Problem termination for the purpose of data analysis was coincidental with the "98 ENTER" command.

Post-test

After finishing the 20 search tasks, subjects were requested to complete the videotex-applications questionnaire and to rate the search method used. They were then debriefed in general terms as to the nature of the research hypotheses.

RESULTS AND DISCUSSION

Overview

With varying degrees of qualification necessary, it can be asserted that most of the research hypotheses formulated prior to this experiment were supported by the empirical outcomes.

Hypotheses 1 and 2

When considered without regard to the nature of the search task, the keyword system demonstrated general superiority in retrieving a greater quantity of relevant information, but failed to produce results which were significantly different from the hierarchical system in other categories of performance. In the number of pages accessed, search time, and time per page, there were no overall differences. However, when search methods were compared in the context of different kinds of search tasks, distinct performance differences materialized in the data. In the solvable unipartite and multipartite problems, the keyword system not only scored more retrieval successes, but also required significantly fewer pages and less time to find this greater quantity of information. In the unsolvable problems, the pattern was reversed. In this instance, it was the hierarchical system which performed better, taking considerably fewer pages and less time to confirm the non-existence of target information.

Hypothesis 3

Both spatial memory and ideational fluency were found to correlate moderately with performance measures. In the hierarchical condition, spatial memory was found to correlate positively with retrieval successes and pages accessed. In the keyword condition, ideational fluency correlated negatively with pages accessed.

Hypotheses 4 and 5

Both search methods were rated positively across performance dimensions, but no appreciable difference was evident between methods. There was an improvement in attitude measures at post-test, but increased favourability was confined to the application actually experienced by subjects and did not generalize to other hypothetical applications. Nevertheless, initial attitudes toward videotex were markedly favourable and were independent of both gender and computer experience.

Search Performance

The data used in all performance-variable analyses were the untransformed mean scores obtained by each subject in each of the three kinds of search task. There were, in other words, three scores per dependent variable for each subject; retrieval success was an exception with two mean scores, for the two kinds of solvable tasks.

Multivariate analysis of variance was used to test for significant differences among performance measures. Because Box's test of the homogeneity of variance-covariance matrices was highly significant, subsequent univariate F -ratio tests involving the within-subjects factor were stringently evaluated using the Greenhouse-Geisser correction to degrees of freedom. Violation of the compound-symmetry assumption also required in the Bonferroni comparisons of pairs of means that the standard-error term used in comparisons involving the within-subjects factor be derived from the difference scores of the pair of conditions concerned.

Multivariate analyses for main effects and interactions in the combined data from solvable and unsolvable problems (Table 11, Appendix H) yielded highly significant F -ratios for Search Method, Search Task, and a Method \times Task interaction.

Univariate tests of Search-Method effects (Table 12, Appendix H) showed the keyword system to have been more successful in finding relevant information than the hierarchy, $F(1,40) = 17.88$, $p < .001$. In the keyword condition, subjects found an average of 1.53 target pages as opposed to 1.12 in the hierarchical condition, or about 37% more relevant information. There were no significant differences in the other performance variables.

The very large F -ratios generated by the univariate tests of Search Task effects confirmed major differences in the demand characteristics

of each type of retrieval problem (Table 13, Appendix H). Even when degrees of freedom were drastically reduced to correct for heterogeneity of variance-covariance, effects across variables were highly significant: for retrieval successes, $F(1,40) = 222.85$, $p < .001$; for pages accessed, $F(1,40) = 46.15$, $p < .001$; for search time, $F(1,40) = 43.39$, $p < .001$; and for time per page, $F(1,40) = 9.81$, $p < .005$.

Comparisons between pairs of Search Task means using the Bonferroni t -test (Table 15, Appendix H) showed that more relevant information was retrieved in the multipartite search tasks (1.89 pages) than in unipartite tasks (.76 pages), $t(40) = 14.30$, $p < .01$. Correspondingly, multipartite tasks consumed 63% more pages (18.35) than unipartite tasks (11.26), $t(40) = 8.64$, $p < .01$, and took almost twice as long to complete (201 seconds versus 115), $t(40) = 10.04$, $p < .01$. While unsolvable tasks consumed, in turn, more pages than multipartite tasks (23.28 pages versus 18.35), $t(40) = 2.82$, $p < .05$, this difference was not matched to any significant extent by longer search times (209 seconds versus 201). Time per page, on the other hand, was significantly lower in unsolvable problems (9.61 seconds) than in multipartite tasks (10.98 seconds), $t(40) = -4.26$, $p < .01$. In sum then, unsolvable problems generally demanded the most effort of searchers and solvable unipartite tasks the least. Solvable multipartite problems fell somewhere in between, demanding a considerable amount of search effort but also returning the largest quantity of relevant information on the extra investment.

Degrees of freedom in the univariate tests for Method x Task interactions (Table 14, Appendix H) were also adjusted by the Greenhouse-Geisser correction. Even so, F -ratios maintained respectable levels of significance: for retrieval successes, $F(1,40) = 5.33$, $p < .05$; for total pages accessed, $F(1,40) = 18.70$, $p < .001$; and for search times, $F(1,40) = 13.13$, $p < .001$.

Subsequent contrasts of pairs of Method x Task means by the Bonferroni procedure (Table 15, Appendix H) were limited to those between Search Methods and within each level of Search Task. As summarized in Table 4 and illustrated in Figure 2, the keyword system scored more hits than the hierarchical system in both the unipartite tasks and the multipartite tasks. Only in the multipartite condition was the difference (2.18 pages versus 1.59 pages) significant, $t(40) = 5.52$, $p < .01$ — a result which suggests a kind of "increasing returns" effect. Means for the total number of pages accessed are shown in Table 5 and plotted in Figure 3. The means for search times appear in Table 6 and Figure 4. As suggested by these data summaries though, differences are probably best discussed in the contexts of solvable and unsolvable problems.

With reference to solvable problems first of all, keyword retrieval was more economical than hierarchical retrieval in both unipartite and multipartite tasks. In unipartite tasks, the keyword system required 34% fewer pages (8.95 versus 13.57) to find target information, $t(80) = 2.60$, $p < .05$. In multipartite tasks, 19% fewer pages were required by

TABLE 4

Target pages successfully retrieved:
means and standard deviations (in parentheses)

Search method	Search task		Marginal means
	Unipartite	Multipartite	
Hierarchy	.64 (.14)	1.59 (.58)	1.12 (.34)
Keyword	.88 (.17)	2.18 (.51)	1.53 (.30)
Marginal means	.76 (.19)	1.89 (.62)	

FIGURE 2

Mean number of retrieval successes as a function of search task.
Vertical bars represent 95% confidence intervals.

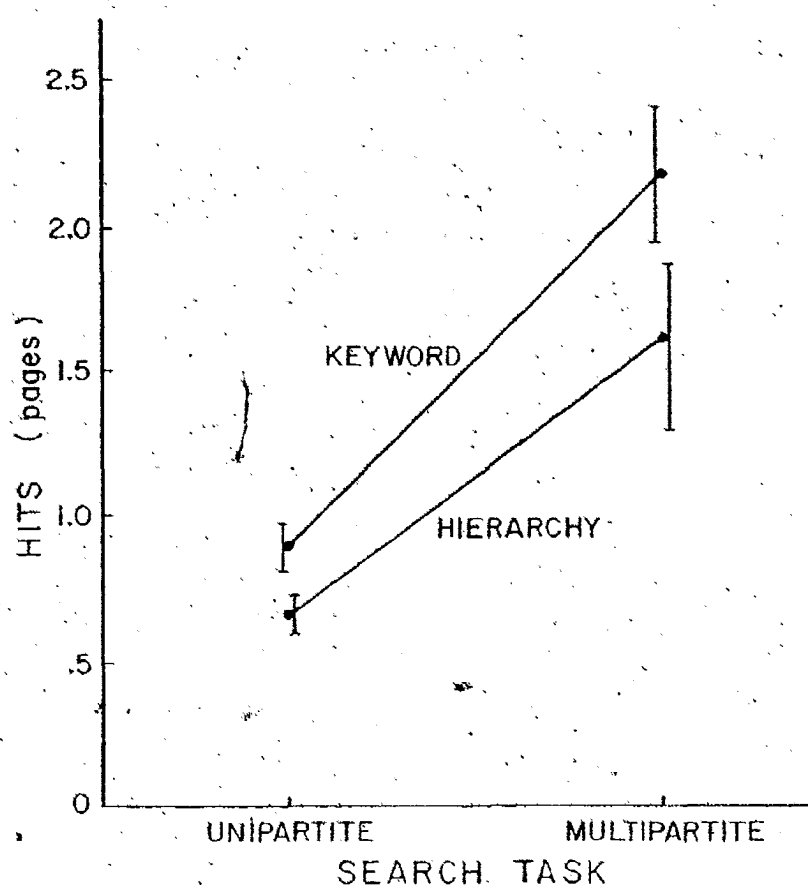


TABLE 5

Total pages accessed:
means and standard deviations (in parentheses)

Search method	Search task			Marginal means
	Unipartite	Multipartite	Unsolvable	
Hierarchy	13.57 (4.01)	20.24 (5.84)	18.73 (7.81)	17.52 (4.34)
Keyword	8.95 (2.94)	16.45 (4.72)	27.82 (10.58)	17.74 (4.65)
Marginal means	11.26 (4.19)	18.35 (5.59)	23.28 (10.27)	

FIGURE 3

Mean number of pages accessed as a function of search task.
Vertical bars represent 95% confidence intervals.

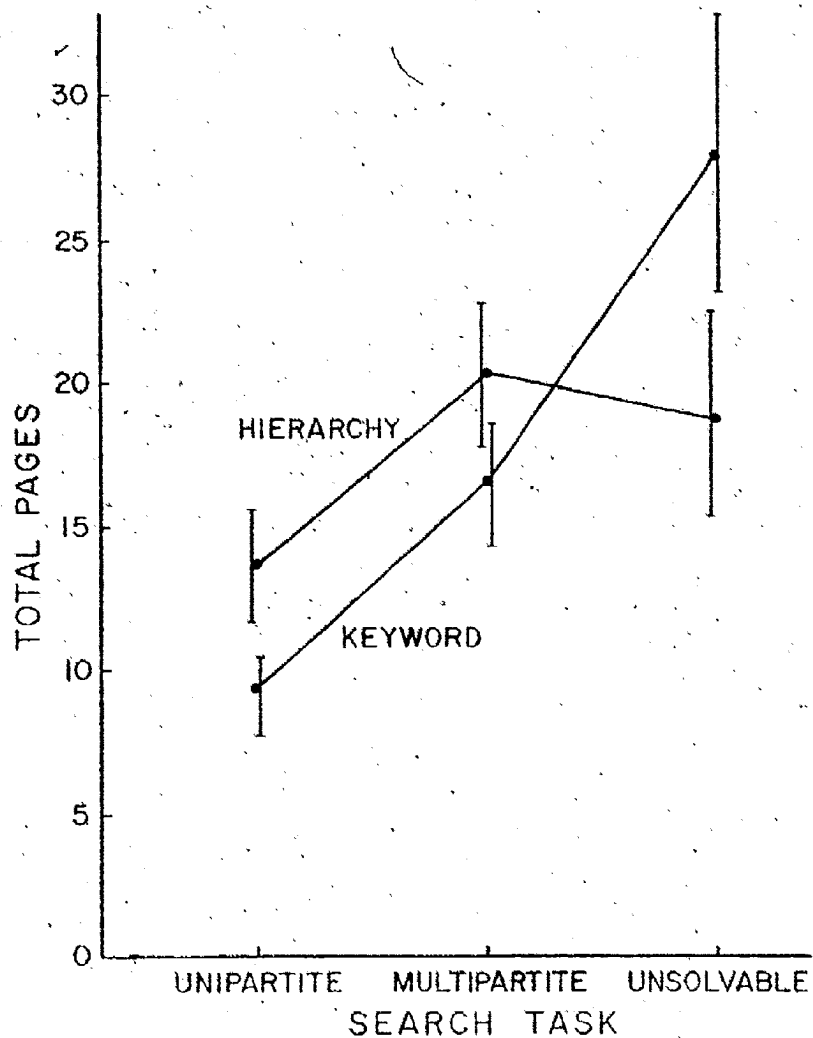


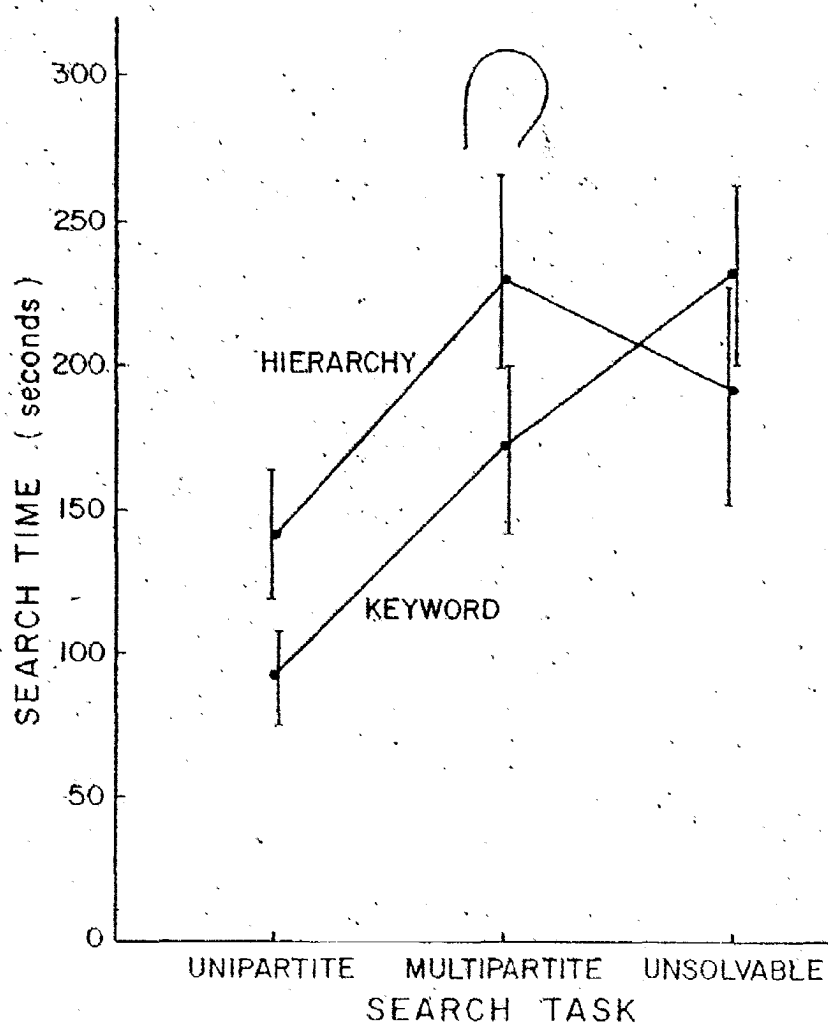
TABLE 6

Search times (seconds):
means and standard deviations (in parentheses)

Search method	Search task			Marginal means
	Unipartite	Multipartite	Unsolvable	
Hierarchy	140.13 (48.85)	230.91 (73.63)	187.46 (79.02)	186.17 (49.61)
Keyword	89.35 (34.31)	170.42 (63.35)	231.05 (67.96)	163.61 (44.35)
Marginal means	114.74 (48.98)	200.66 (74.43)	209.26 (76.06)	

FIGURE 4

Mean search times as a function of search task.
Vertical bars represent 95% confidence intervals.



keyword retrieval (16.45 versus 20.24); but this difference was not statistically significant. The keyword system was also faster than the hierarchical: 51 seconds faster in the unipartite tasks (89 seconds versus 140), $t(80) = 3.20$, $p < .01$; and 61 seconds faster in the multipartite tasks (170 seconds versus 231), $t(80) = 3.81$, $p < .01$.

When unsolvable problems were considered independently of the others, the relative standings of the two search methods underwent a dramatic turnabout. In these tasks, the hierarchical searches used up 33% fewer pages than keyword searches (18.73 versus 27.82), $t(80) = -5.11$, $p < .01$, and were completed about 44 seconds faster on the average (187 seconds versus 231), $t(80) = -2.75$, $p < .05$.

While the overall F -test did not show any significant differences in time per page due to Method \times Task effects, Table 7 and Figure 5 reveal a pattern of relationships that is sufficiently regular to justify a few comments. As the search task became more complex, unipartite to multipartite, the average time per page increased from 10.22 seconds to 10.98 seconds. As the search task consumed even more pages in the unsolvable problems, not only did the average time per page drop in both the hierarchical and keyword conditions (to 10.45 and 8.77 seconds respectively), but the speed gap between search methods widened.

When solvable problems were subjected to a separate analysis, multivariate tests of significance (Table 16, Appendix H) again showed main effects for Search Method [$F(4,37) = 10.35$, $p < .001$] and Search

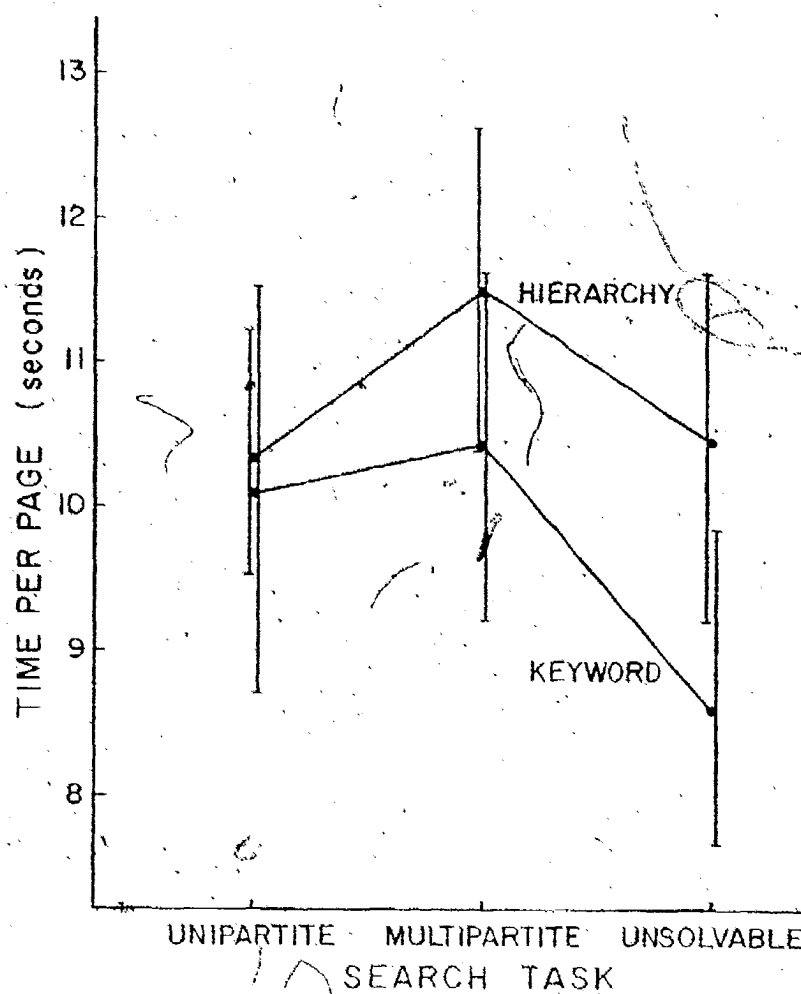
TABLE 7

Time per page (seconds):
means and standard deviations (in parentheses)

Search method	Search task			Marginal means
	Unipartite	Multipartite	Unsolvable	
Hierarchy	10.34 (1.79)	11.52 (2.50)	10.45 (2.67)	10.77 (2.06)
Keyword	10.09 (3.03)	10.44 (2.67)	8.77 (2.24)	9.77 (2.39)
Marginal means	10.22 (2.46)	10.98 (2.61)	9.61 (2.58)	

FIGURE 5

Mean time per page as a function of search task.
Vertical bars represent 95% confidence intervals.



Task [$F(4,37) = 66.51, p < .001$], but did not indicate any interaction. Subsequent univariate tests for Search-Method effects (Table 17, Appendix H) generally confirmed the interaction results observed in the combined data, namely, showing the keyword system superior on three out of four performance measures. As previously noted, the keyword system retrieved more relevant information and was more efficient in so doing. On the average, keyword searches in solvable problems required 25% fewer pages (12.70 versus 16.91), $F(1,40) = 14.22, p < .01$, and were 30% faster (130 seconds versus 186), $F(1,40) = 13.17, p < .01$. Differences due to Search-Task characteristics (Table 18, Appendix H) were identical with those reported above except that F -ratios tended to be larger.

As an aid in identifying the most distinguishing characteristics of hierarchical and keyword retrieval and as a test of the redundancy of performance measures, both two-group (Search Method) and six-group (Method x Task) stepwise discriminant analyses were carried out on the total data set. In both cases, the relative discriminating power of the variables, as reflected in their order of extraction, was identical: retrieval success followed by pages accessed and search time. In the two-group analysis, no one variable had by itself significant discriminating power, so that loadings on the single resulting function (Table 8) did not reflect the order of extraction.

As the classification results indicate, the derived linear composites were not particularly accurate in differentiating group membership: 66.7% of the cases correctly classified in the two-group

TABLE 8

Standardized coefficients and classification results
for discriminant models

Analysis	Canonical correlation	Classification accuracy	Variables included	Standardized coefficients
<u>Two-group</u>				
Function 1	.34	66.7%	Hits	.71
			Pages	1.46
			Search time	-1.60
<u>Six-group</u>				
Function 1	.93	72.2%	Hits	1.00
			Pages	-.28
			Search time	.01
Function 2	.66		Hits	.11
			Pages	.64
			Search time	.41

analysis and 72.2% in the six-group analysis. Nevertheless, the six-group analysis did produce some useful information. The first discriminant function was heavily loaded with the recall measure, hits, while the second function was loaded with measures of retrieval effort, pages accessed and search time. Thus, from the scatterplot in Figure 6, it can be seen that groups in, or closest to, the lower right quadrant (keyword-unipartite and keyword-multipartite) represent the most recall with the least effort, while the group in the upper left quadrant (keyword-unsolvable) represents the least recall with the most effort. This illustrates, quite literally, the combined effects of the Method \times Task interaction described above.

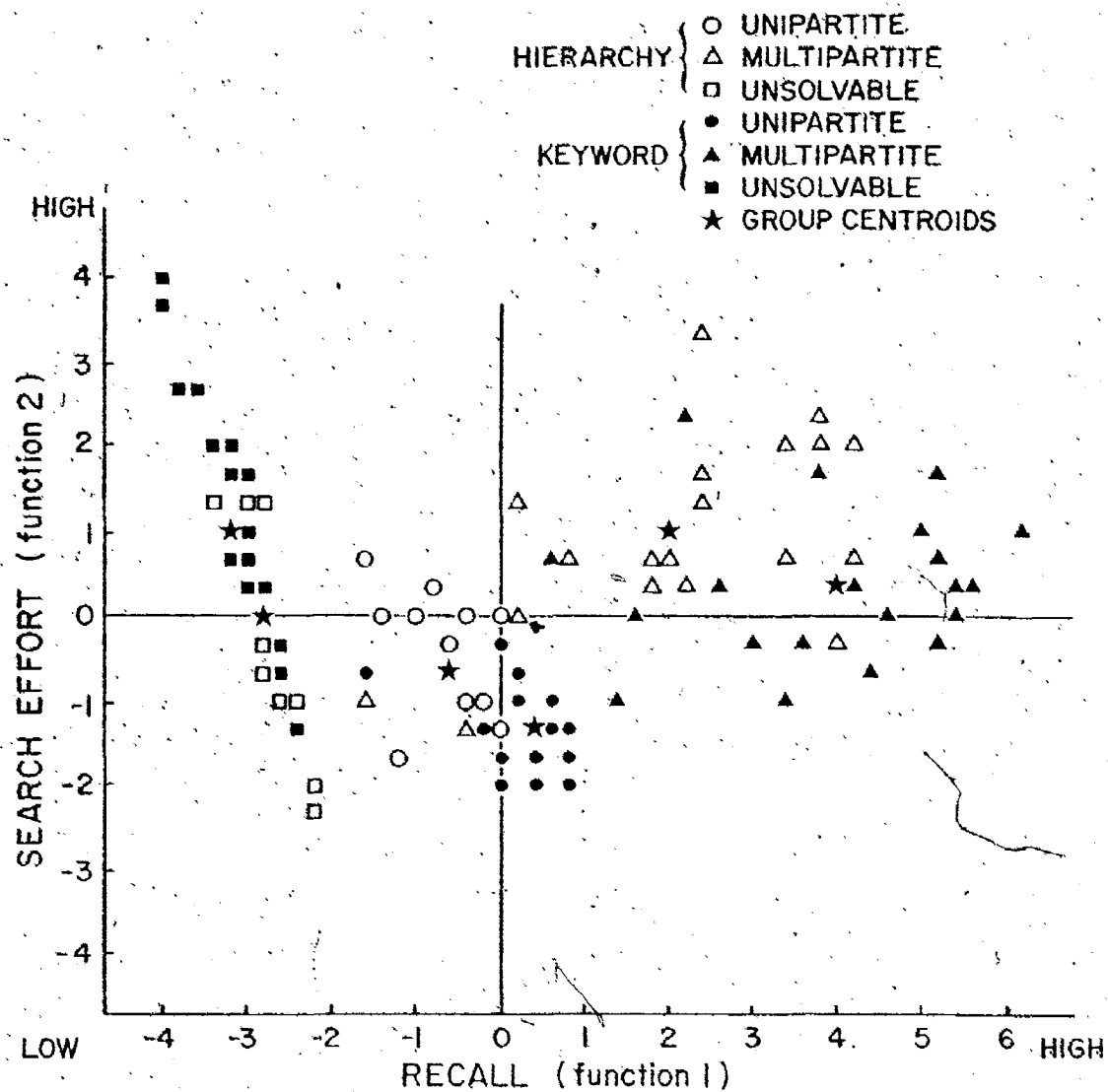
Simple correlations among dependent measures showed that the number of pages accessed was positively correlated with search time in both search conditions (hierarchical $r = +.73$; keyword $r = +.58$). Notably, these associations were substantially weaker than the correlation of $+.90$ found in pilot research. The only other sizeable correlation was between pages accessed and retrieval success ($r = +.60$), this in the hierarchical data set.

Hierarchical Retrieval or Keyword Retrieval?

As the results have shown, there are real and substantial differences between hierarchical retrieval and keyword retrieval in videotex. Specifically, the keyword system demonstrated superior performance in solvable search tasks while, contrary to expectations,

FIGURE 6

Scatterplot resulting from discriminant analysis of performance data grouped according to Search Method and Search Task. The horizontal axis represents Canonical Discriminant Function 1, which is heavily loaded with the measure of recall, information hits (see Table 8). The vertical axis represents Canonical Discriminant Function 2, which is heavily loaded with measures of search effort, pages accessed and search time. Observations in the lower right quadrant are indicative of high recall with little work, and thus represent the most effective combinations of search method and retrieval task. Observations in the upper left quadrant signify low recall and much work, and represent the least effective combinations.

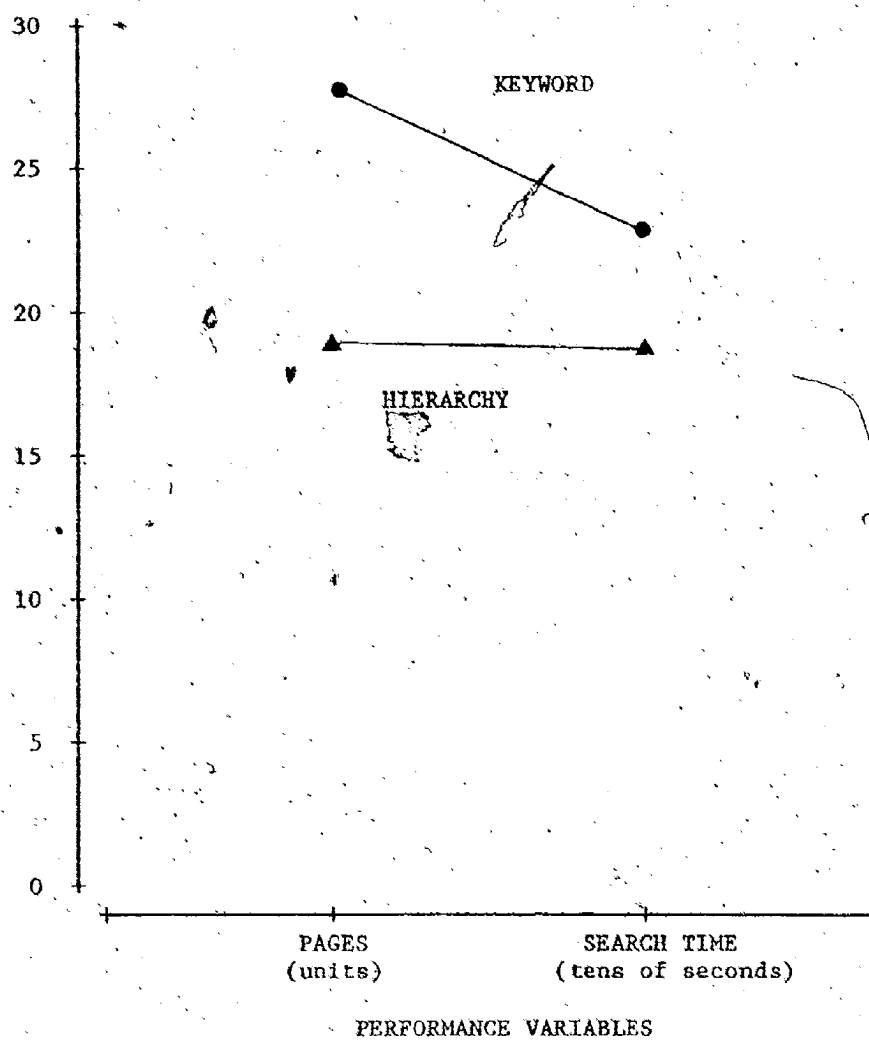


the hierarchical performed more efficiently when target information was not actually present in the database. Although it might be tempting to argue that the overall trend favours keyword retrieval and that the results in the unsolvable problems might be chance anomalies, the differences in the latter case are too large and too consistent to support such a contention (see Figure 7). There are, in retrospect moreover, sound reasons for the hierarchy's better showing in unsolvable tasks.

If we accept the book model of indexing put forward by Thomas and Schabas (1981), according to which the hierarchical index functions as a table of contents and the keyword directory as a back-of-the-book index, then one chief advantage of the hierarchy becomes apparent. A reading of the menu of information categories and category descriptors at the top of the hierarchy quickly establishes for the user rough guidelines as to the size, shape, boundaries, and content domains of the database. In the keyword directory on the other hand, this kind of meta-information does not seem to come as easily. The keyword directory does not, after all, have any kind of gross structure; it is, column after column, a parade of detail. Thus, at the beginning of any search, the hierarchy user can logically eliminate from consideration those portions of the database which are unlikely to contain the target information and, concurrently, limit his/her exploration to the most promising categories. Naturally, disconfirmation in unsolvable search tasks will be relatively fast when the range of the search is so reduced. The keyword user, meanwhile, is always obliged to deal with

FIGURE 7

Summary comparison of search-method performance characteristics in unsolvable problems



the entire database, the boundaries of which are probably unknown or else somewhat fuzzy, and, if unsuccessful with the first search term, must try some variant or new term until, having exhausted his/her patience or ingenuity, the chase is given up. But even after extensive searching, the keyword user still cannot be entirely confident that the deficit is in the database and not in his/her ability to generate the right keyword.

Since the keyword directory runs a distant second to hierarchical indexing in unsolvable problems, it seems appropriate to ask if keyword retrieval provides offsetting benefits in other kinds of search tasks. Upon comparing the performance of the two search methods in solvable problems, it may have been noted that the differences were not symmetrical from unipartite to multipartite tasks. Only in search time did keyword retrieval prove to be consistently better. Otherwise, in unipartite tasks, the keyword system retrieved about the same amount of information in significantly fewer pages while, in multipartite tasks, it retrieved significantly more information in about the same number of pages. Although the graphical representations of these relationships (Figures 8 and 9) show a convincingly consistent performance spread, it would be more satisfying if the numerical pattern could be made as tidy. One way to do this is to calculate the cost in system pages and time for a unit item of information. The resulting ratios show that, in unipartite tasks, one page of information obtained by keyword retrieval costs 10 pages or 191 seconds and, by hierarchical retrieval, 21 pages or 218 seconds. In multipartite tasks, one page of information

FIGURE 8

Summary comparison of search-method performance characteristics in unipartite problems

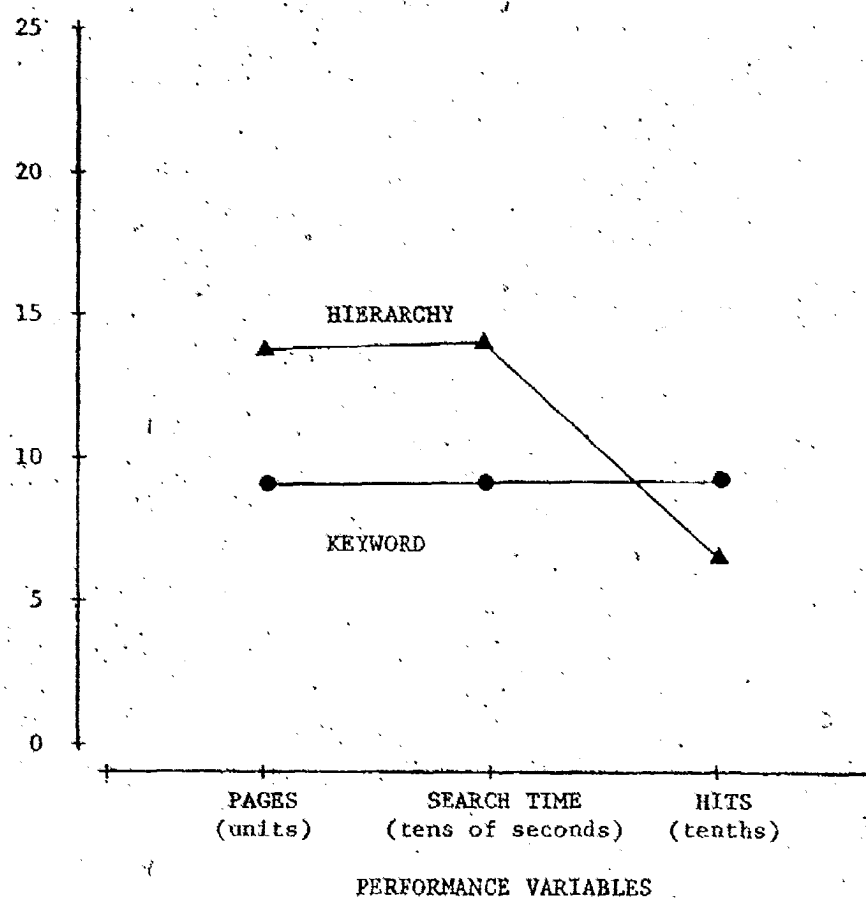
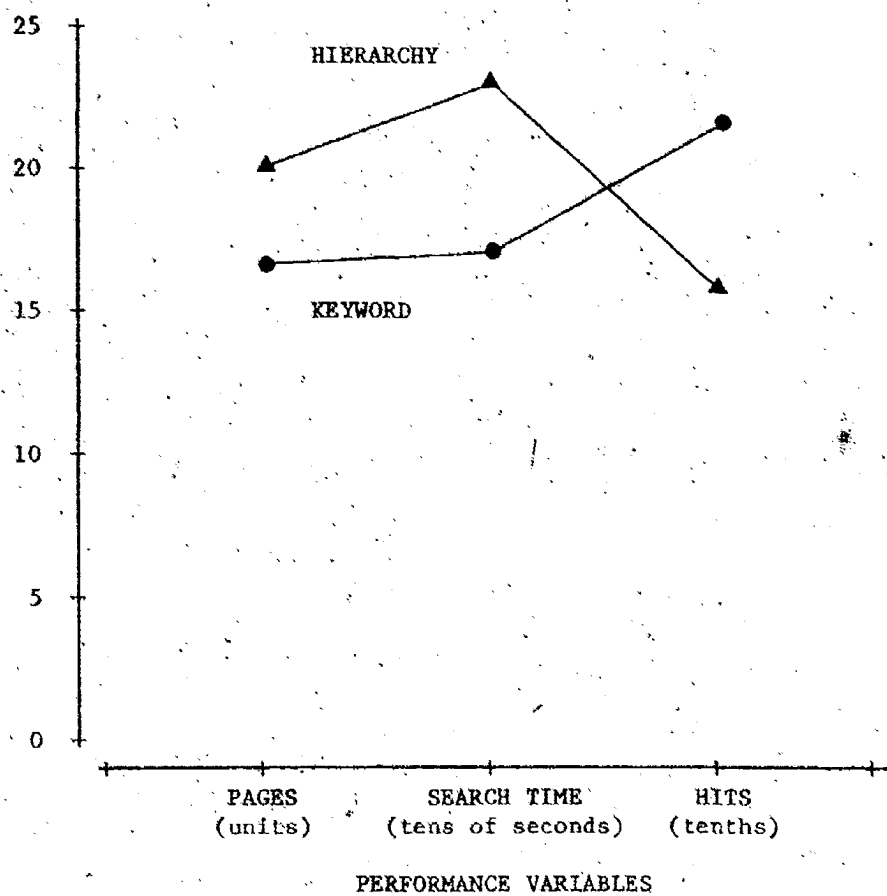


FIGURE 9

Summary comparison of search-method performance characteristics in multipartite problems



retrieved by the keyword index costs 7.5 pages or 78 seconds and, by the hierarchical index, 12.7 pages or 145 seconds. A simple and consistent relationship is immediately evident. To retrieve a given quantity of information, regardless of the type of task, it takes half as many pages and half as long by keyword as it does by hierarchical retrieval. This is equivalent to a 50% improvement in search efficiency, which is more than adequate compensation for the weaknesses noted in the unsolvable problems. An incidental point of note is that the cost for an item of information, in pages accessed or search time, can be reduced by $1/4$ to $1/3$, regardless of search procedure, if more than one item is targeted during each search.

Information-Processing Demands

Another facet of search performance which merits some attempt at an explanation concerns the finding that the average time spent per page, in both search conditions, increased as the search task became more complex but declined significantly in long, unsolvable tasks. This is logically sensible. As a problem increases in dimensionality, more information processing by the searcher is necessary on each page; there are, quite simply, more options to evaluate. Thus multipartite tasks show an increment in time per page over unipartite tasks. But, after extensive and unsuccessful searching, as the suspicion gets stronger that the target information may not in fact be in the database, the mode of information processing probably shifts from relatively slow, careful scanning to very rapid scanning of the video display. Consequently,

unsolvable tasks show a decrement in time per page when compared to multipartite tasks.

Besides these information-processing similarities between search methods, there appear to be qualitative differences as well. For instance, the difference in time per page between search methods rose from .25 seconds to about 1 second to almost 2 seconds as the task required more pages and time. While this between-groups difference was not statistically significant, presumably because of low power, it did have practical consequences in the unsolvable tasks; although keyword retrieval took some 50% more pages, this large page differential was offset in the search-time differential by the ability of keyword searchers to traverse the extra pages at a faster rate. Learning effects can be ruled out as an explanation since the order of problems was completely randomized for each subject. Why this difference appeared probably has something to do with the unique information-processing demands of each search method. In the hierarchical procedure, the searcher begins with a concept and scans successive menus of randomly ordered terms looking for a semantic match. In the keyword procedure, by way of contrast, the searcher generates a term and scans successive menus of alphabetically ordered terms looking for a typographical match. Thus it seems reasonable to assume that, on indexing pages, there will be a small time differential between scanning/interpreting/comparing in an unstructured list of category terms and simple scanning/comparing in a structured list of keywords. But in long searches, which typically increase the number of indexing

pages accessed, the cumulative effect of this information-processing difference will be detectable as a differential in time per page; hence the apparent spread between search methods in multipartite problems and unsolvable problems.

Search Performance and Subject Characteristics

Performance variables which survived stepwise discriminant analysis were grouped with measures of subject characteristics in a series of canonical correlations and stepwise multiple regressions for the purposes of determining whether any relationship existed between select subject characteristics and the extent to which performance could be predicted by subject measures. In these analyses, performance measures were weighted to reflect the proportional distribution of the three kinds of search tasks in the original test battery.

Canonical-correlation analyses were performed on the total, hierarchical, and keyword data sets of performance measures and the subject characteristics of gender, age, level of computer experience, spatial-memory ability, ideational fluency, and handedness. No significant or meaningful canonical variables emerged from these analyses.

In the stepwise multiple regressions, each major performance variable was paired in turn with the subject variables. Regressions on the total data set (Table 9) identified spatial memory as a significant

TABLE 9

Significant correlations
between subject variables and performance variables

Data set	Predictor	Criterion	Pearson r	R^2 (%)	Adj R^2 (%)
Total	Spatial memory	Hits	.42**	17.8	15.8
	Ideational fluency	Pages	-.31*	9.8	7.6
Hierarchy	Spatial memory	Hits	.58**	34.0	30.5
	Spatial memory	Pages	.44*	19.7	15.5
Keyword	Ideational fluency	Pages	-.54*	29.5	25.8

* $p < .05$

** $p < .01$

positive correlate of retrieval success, or recall [$r(40) = +.42, p < .01$], and ideational fluency as a significant negative correlate of pages accessed, or work [$r(40) = -.31, p < .05$]. When regressions were subsequently run using the data set from each search method, the resulting coefficients showed stronger associations. In these associations, furthermore, each ability was tied to a particular search method. As had been predicted, spatial memory scores correlated with hierarchical performance: $+.58$ with retrieval success ($p < .01$) and $+.44$ with the number of pages accessed ($p < .05$). In the keyword condition, by way of contrast, it was ideational fluency which had been predicted as the more likely mediator of search performance, and scores correlated $-.54$ with the number of pages accessed ($p < .05$). Details of the regression analyses are shown in Tables 19-22, Appendix H.

Compatibility of Users and Retrieval Systems

Can anyone use either retrieval system with relatively equal effectiveness and efficiency? Probably not. All things not being equal from user to user, the results of the regression analyses suggest that search performance can be either enhanced or impaired depending on the particular combination of retrieval method and user abilities.

In hierarchical searches, about 30% of the variability in retrieval success and 15% of the variability in pages accessed can be accounted for by a spatial-memory factor. Thus, effectiveness, or recall, seems to depend in part on the user's ability to acquire an appropriate

spatial model, or map, of the information hierarchy and subsequently apply that knowledge in navigating the actual territory. It is surprising though to find the same ability positively correlated, even if weakly, with an index of retrieval effort; the relationship one would expect in such a case would be a negative one. But, in the hierarchical case, retrieval success was positively correlated ($r = +.60$) with pages accessed. This unavoidable necessity to work harder for one's rewards in the hierarchical system could be taken as a testament to its inherent rigidity and intractability.

In keyword searches, about 25% of the variability in pages accessed was accounted for by ideational fluency. The negative correlation simply means that the relationship is what one would expect; the more fluent the searcher as a rule, the less work necessary. It does not follow, however, that a searcher with a conceptually richer vocabulary or one who is given to free-associating will find more relevant information. Fortunately for the user, neither effort nor fluency appear to be tied to recall in the keyword system. All that can be said with any confidence is that a fluent individual might be slightly more efficient at the task.

Ratings of Search Methods

Subject ratings of the hierarchical retrieval system were favourable on all ten dimensions but showed little variability from one dimension to another, ranging from an absolute low of 4.3 to an absolute

high of 5.9 across all 7-point scales. Ratings of the keyword system were much the same, ranging from 4.7 to 6.0. To determine whether differences in performance between search methods had been translated into differences in subjective evaluations, rating data were tested by Hotelling's T^2 statistic. The test showed no differences attributable to the Search Method factor (Table 24, Appendix H) and so plans for more detailed analyses had to be abandoned. Otherwise, the only item of note to be extracted from the raw scores was that subjects found the computerized version of the academic calendar more interesting and more efficient than the print version.

Subject Perceptions of the Search Methods

The fact that large differences in objective performance measures were not reflected in detectable differences in subjective ratings on related performance dimensions raises the possibility that there are in reality no psychologically meaningful differences between search methods. However, the reactions of subjects who were shown both retrieval systems after the experiment contradict this view. The few students who saw both systems in action were considerably more impressed by keyword retrieval, and several indicated that, had they tried both, their ratings would have been quite different.

Attitudes toward Videotex

McNemar's χ^2 for correlated measures was used to test for significant attitude change as the result of exposure to a videotex application. To determine whether pre-exposure attitudes to videotex applications were linked to the subject variables of gender and computer experience, χ^2 tests of independence were carried out.

Despite the fact that only two of the 42 subjects had previously encountered commercial videotex and that slightly less than half were computer-experienced, pre-exposure attitudes to several hypothetical videotex applications (see Appendix E) were decidedly favourable (Table 10). Only the application actually experienced during the experiment, a computerized calendar-information system, produced a significant increase (28.6%) in the proportion of favourable ratings ($\chi^2 = 10.08$, $p < .01$). Other applications showed marginal, non-significant gains. Tests of possible pre-disposing influences on initial attitudes, i.e. computer experience and gender, did not provide evidence of any connection (Table 25, Appendix H).

Subject Perceptions of Videotex

While a large majority of the student subjects were initially very receptive to several hypothetical videotex applications, only in the application actually experienced was there a significant increase in the expression of favourability. It seems almost a contradiction then;

TABLE 10
Attitudes toward videotex applications

Application	Percentage in favour		χ^2	Binomial probability ^a
	Pretest	Post-test		
Computerized calendar	69.0	97.6	10.08*	—
Calendar + transactions	81.0	88.1	—	.38
Videotex calendar service	66.7	71.4	—	.69
Commercial videotex	81.0	85.7	—	.63

^a The binomial distribution was used to calculate probabilities when the number of changes was less than ten.

* $p < .01$

to state in the same breath, that prior computer experience was not linked to pretest favourability. One can speculate of course that pretest measures may actually have measured a non-specific pro-computer attitude. There is some evidence to support this view. Market trials (Eissler, 1981; OCLC, 1981) have shown that it is chiefly among the young, well educated members of the middle class that computer consciousness has been most enthusiastically embraced. In a sample of university students, a similar outlook should not be surprising. Such being the case among the students who initially indicated favourable attitudes, general cultural trends and not personal experience may have pre-determined their responses, whereas for those who were initially non-committal, or who held unfavourable opinions, the personal experience of using a videotex system may have played the stronger role in winning qualified support.

GENERALIZABILITY AND IMPLICATIONS OF EXPERIMENTAL FINDINGS

Generalizability to Videotex

The structure and operating characteristics of the experimental information system were modelled on those features common to videotex systems such as Prestel (Rimmer, 1979; Dew, 1980; Stewart, Cox & partners, 1980; Williams, 1981), CAPTAIN (Inoue, 1980; Harashina & Kitamura, 1981), and Telfidon (Parkhill, 1980; Phillips, 1980; Ball, 1981; Latremouille & Lee, 1981; Dillon & Tombaugh, 1982). Three differences should be noted however. Firstly, no attempt was made to emulate either of the currently available videotex graphics systems; the experimental database was text only. Secondly, the data transmission rate of the host computer (9600 baud) was considerably higher than what can be obtained over a conventional telephone line (300-1200 baud); nevertheless, this feature was an experimental constant and was judged unlikely to contribute to the real or perceived advantage of either search method. Thirdly, because computer terminals, rather than television monitors, were used as display devices the maximum display capacity of the experimental system was 24 lines of 80 columns (1920 characters) in contrast to the 20 lines of 40 columns (800 characters) available on a television screen; again, this feature was an experimental constant and was not expected to invalidate any comparisons. Furthermore, the maximum number of keywords and document titles listed on the keyword indexing pages was limited to 16, a number which can easily be accommodated by

commercial videotex.

In videotex systems such as Prestel, CAPTAIN, and Telidon, the number of choices available on any menu page is limited to a maximum of nine or ten, in part by the particular page-numbering convention adopted to identify nodes, or levels, in the hierarchy (Inoue, 1980; Winsbury, 1980; Ball, 1981), and also, perhaps in deference to Miller's magic number seven-plus-or-minus-two (Miller, 1967), by a desire to keep the shape of the hierarchy as close to the optimum as possible, that is, narrow and deep rather than broad and shallow (Rimmer, 1979; Lee, 1980). In the hierarchy developed for this research, these principles were generally observed. It was six categories wide at the top and varied between four and seven levels in depth. However, certain characteristics of the experimental database necessitated the creation of some menu pages with considerably more than ten choices. In such situations (the list of Academic Departments for example), category terms were usually listed alphabetically, a simple adaptation which has been demonstrated not only to counteract the adverse consequences of violating Miller's law but actually, in some circumstances, to improve search performance (Stewart, Cox & partners, 1980). Additionally, to lend some flexibility to what is characteristically a rigid structure, topically related documents and pages were cross-referenced within information documents where appropriate.

In principle, fairly sophisticated keyword systems which allow

for keyword entry at a keyboard, post-coordination of terms, and customized response frames, could be developed for videotex systems. However, two of the consequences of relying on a numerical keypad for input, and its requisite menus, were that entry terms had to be pre-coordinated, and response pages had to be fixed in order to preserve a retracing capability, thus resulting in an operating procedure seemingly at cross-purposes with the aims of the indexing system. To be explicit, the savings in search time and pages accessed promised by a keyword approach could have been squandered on the extra pages required for menus. Nonetheless, the number of levels optimally required to access any document was reduced to a maximum of five and in most cases did not exceed four. In testing, moreover, the practical merits of this approach were amply demonstrated.

Implications of Findings

Some of the results generated by this experiment point to relatively straightforward lessons for the conduct of information-retrieval research. For example, the failure to obtain significant rating differences between search methods clearly indicates that subjects require an experiential basis in order to make valid comparisons. This argues strongly, of course, for a within-subjects design whenever subjective comparisons of retrieval systems are at issue. Secondly, the fact that stepwise discriminant analyses did not partial out any of the performance variables entered (*i.e.* retrieval successes, pages accessed, search time) demonstrates

that each of these indices taps a unique source of variability, even though there is significant overlap as well. Consequently, all of these measures at least can be used as dependent variables without fear of redundancy. Finally, the strong interaction between search method and search task underscores an absolute requirement to sample a variety of search tasks in information-retrieval research.

Undoubtedly, the conclusions drawn about the relative effectiveness of either search method would have been quite different had unsolvable tasks not been included in the design. Similarly, it might not have been obvious that multipartite searches are more economical than a string of unipartite searches if both had not been investigated.

The import of the videotex attitude measures is less certain, but, if there are any messages in these results for the marketing of videotex, they would seem to be two: that young adults with some university or college are primed for computerization; and that, among those who regard videotex unfavourably or skeptically, free trial services might be successful in attracting interest. Favourable attitudes and high interest do not entail, however, a willingness to pay for videotex services. Perceived utility, cost, and other issues would presumably be more influential in any decision to subscribe, and these factors could best be identified by the kind of large-scale trials now under way.

Turning our attention to search performance, we encounter a number of results which have major implications for the design and

operation of information-retrieval systems. First of all, the direct and indirect evidence for a spatial factor mediating performance in computer systems (Barif & Lusk, 1977; Benbasat, Dexter & Masulius, 1981; Egan, Bowers & Gomez, 1982; Gomez & Egan, 1982; Robertson, McCracken & Newell, 1981) receives additional support from the association demonstrated between spatial-memory scores and measures of recall and efficiency in the hierarchical search condition. As for hierarchical systems in particular, the confirmation of a relationship between spatial memory and retrieval performance gives a few more teeth to the proposals made by Lochovsky and Tsichritzis (1981) for including real-time navigational aids in such systems.

While search performance in the keyword system was not similarly dependent on a navigational factor, the poor results obtained in unsolvable problems clearly imply, when compared to the results of hierarchical searches, that something was missing. Thus, in spite of the keyword directory's superiority in handling solvable problems, it must be concluded that this kind of keyword system, and perhaps any kind of keyword system, cannot stand alone as the one and only indexing device. A hierarchical index, or some other database map, must be available, if for no other reason than to lend the undisciplined particulars of the keyword directory some semblance of topical structure.

The strong interaction between search methods just noted and the performance differences between kinds of search tasks also say

something about the heuristics of search strategy when both hierarchical and keyword-directory retrieval systems are available to a user. As a first principle, users should avoid looking for information one item at a time; it will probably be cheaper to do all of one's electronic shopping in one database trip. Secondly, if the user is reasonably sure that the target information is in the database, then the search method of choice should be the keyword system; otherwise, if there is strong doubt, then a hierarchical search should be considered from the outset or after failing on at least two attempts by keyword. All other things being equal, the complementary strengths of both systems can be exploited in this way.

Even when the differential effects of spatial and fluency abilities are entered into the equation, these principles remain intact. Given that keyword retrieval is, to begin with, more efficient than hierarchical, and that recall is independent of either spatial memory or ideational fluency in the keyword system but is correlated with spatial ability in the hierarchical system, keyword retrieval comes out heavily favoured as a first choice, even if the user has a concept-poor vocabulary. Further research in this area is warranted, however, since part of the success of keyword retrieval may be attributable to the browsing facility which was afforded by a visible keyword directory. In most keyword systems, the authority list, or thesaurus, is invisible to the searcher.

It cannot have escaped notice that the additional power and

flexibility purchased by the keyword-directory approach come at a considerable cost in system space. Whereas indexing pages in the hierarchical index numbered about 60 to 70, the indexing pages for the keyword system numbered 434, of which 66 (15%) were actual keyword pages, and the remaining 368 (85%), document-title pages. This experimental database was quite modest, about 700 information pages. Extrapolating indexing-page requirements to large databases does not encourage optimism. One could argue, of course, that the cost of mass storage is falling steadily, so that 10,000 or even 100,000 indexing pages would not be too many in a large system. Regardless of the merits or flaws in this reasoning, it should be possible to reduce the number of indexing pages required for keyword retrieval to a fractional amount. Document-title pages, the big hoarders of computer memory, could be dropped altogether as a fixed feature and synthesized in real time as the need arises. Selection of a given keyword would open a master-file of keywords, each paired with the document titles to which it had been referenced. From this list of document titles and corresponding page numbers, the computer would create one or more document-title pages using dummy pages designated for this purpose. This solution would limit keyword indexing pages in the system to those containing keyword lists, thereby holding the number of stored pages at a level competitive with any hierarchical index. Some deterioration of system response time could be an unavoidable concomitant of real-time page creation, but experimentation with the parameters implicated should readily demonstrate whether or not the effects are psychologically important.

As a final comment, it must be stated that the tacit assumption underlying the discussion of the keyword and hierarchical retrieval results has been that the idiosyncratic operating demands and characteristics of the two systems were responsible for the performance differences observed. This is by no means entirely justifiable. It was pointedly remarked in the report of pilot research that subjects differed in their ability to catch on to the optimum strategy for keyword retrieval, and that experimental instructions were revised accordingly to give subjects the benefits of coaching and system modelling as well as whatever advantages might naturally accrue from objective differences between the systems. Whether this manipulation contributed to performance differences, and how large the effect was, are issues which clamour for investigation. If the effects of coaching users in search strategy and of modelling the retrieval system turn out to be substantial, then these findings would have consequences for user instruction in general and might also shed some light on the failure of comparisons between hierarchies and alphabetical directories (Stewart, Cox & partners, 1980; Tombaugh & McEwen, 1982) to find significant performance differences. In any event, this experiment has shown that, when coaching and system modelling are provided, not only is information retrieval by keyword directory feasible but such a system is easy to use and works extremely well.

REFERENCES

- Aysan, M. Project Ida: home of the future, In Richard Larratt (ed.), Inside Videotex: The future . . . now. Proceedings of a seminar held 13-14 March 1980; Toronto: Infomart, 1980, 60-75.
- Ball, A. J. S. Videotex: chimera or dream machine, Canadian Library Journal, 1981, 38, 11-15.
- Barif, M. L. & Lusk, E. J. Cognitive and personality tests for the design of management information systems, Management Science, 1977, 23, 820-29.
- Benbasat, I.; Dexter, A. S. & Masulius, P. S. An experimental study of the human/computer interface, Communications of the ACM, 1981, 24, 752-62.
- Canadian Broadcasting Corporation, "The Nature of Things", Television air date 20 October 1982.
- Card, S. K.; Moran, T. P. & Newell, A. The keystroke-level model for user performance time with interactive systems, Communications of the ACM, 1980, 23, 396-410.
- Cuff, R. N. On casual users, International Journal of Man-Machine Studies, 1980, 12, 163-87.
- Dew, B. An introduction to Prestel, Physics Education, 1980, 15, 15-20.
- Dillon, R. F. & Tombaugh, J. W. Psychological research on videotex, Behavior Research Methods and Instrumentation, 1982, 14, 191-97.
- Dillon, R. F. & Tombaugh, J. W. Providing HELP to Telidon users, In R. F. Dillon (Chair) Psychological Research in the Development of Telidon. Symposium presented at the meeting of the Canadian Psychological Association, Toronto, 1981.
- Doktor, R. Cognitive style and the use of computers and management information systems, Management Datamatics, 1976, 5, 83-88.
- Doyle, L. B. Information Retrieval and Processing, Los Angeles: Melville Publishing Company, 1975.
- Egan, D.; Bowers, C. & Gomez, L. Learner characteristics that predict success in using a text-editor tutorial, In Proceedings of Human Factors in Computer Systems, Gaithersburg, Md., 15-17 March 1982.

- Egan, D. & Gomez, L. Characteristics of people who can learn to use computer text editors, In Proceedings of the ASIS Annual Meeting, 19, Columbus, Ohio, 17-21 October 1982.
- Eissler, C. O. Market testing video-text: Oak's Miami teletext system, In Proceedings, Videotex '81 International Conference & Exhibition, 20-22 May 1981; Toronto: Infomart, 1981, 57-70.
- Ekstrom, R. B.; French, J. W.; Harman H. H. & Dermen, D. Manual for the Kit of Factor-Referenced Cognitive Tests 1976, Princeton, New Jersey: Educational Testing Service, 1976.
- Feeley, J. The Canadian Telidon field trials, In Proceedings, Videotex '81 International Conference & Exhibition, 20-22 May 1981; Toronto: Infomart, 1981, 17-25.
- Forgie, D. Videotex research and development: the Canadian context and contribution, Canadian Journal of Information Science, 1981, 6, 25-34.
- Frankhuizen, J. L. & Vrins, T. G. M. Human factors studies with viewdata, In Proceedings of the Ninth International Symposium on Human Factors in Telecommunications, Red Bank, New Jersey, 29 Sep - 3 Oct 1980; 1-7.
- Fraser, J. From "POTS" to "PANS": videotex development in Canada, In Viewdata and Videotext 1980-81: A Worldwide Report; White Plains, New York: Knowledge Industry Publications, 1980, 603-12.
- Gaujard, P. Videotex: immediate prospects, In Richard Larratt (ed.) Inside Videotex: The future . . . now, Proceedings of a seminar held 13-14 March 1980; Toronto: Infomart, 1980, 104-09.
- Godfrey, D. All information in all places at all times, In David Godfrey & Douglas Parkhill (eds.) Gutenberg Two, 2nd ed., Toronto: Press Porcupine, 1980, 1-11.
- Gomez, L. & Egan, D. Age, spatial memory, and computer text editing, Paper presented at Psychonomic Society Conference, Minneapolis, November 1982.
- Guilford, J. P. Creativity and learning, In Donald B. Lindsley & Arthur A. Lumsdaine (eds.) Brain Function IV, Berkeley: University of California Press, 1967, 307-26.
- Hakstian, A. R. & Cattell, R. B. Manual for the Comprehensive Ability Battery, Champaign, Illinois: Institute for Personality and Ability Testing, 1976.
- Harashima, S. & Kitamura, T. Current status of the CAPTAIN system, In Proceedings, Videotex '81 International Conference & Exhibition, 20-22 May 1981; Toronto: Infomart, 1981, 113-21.

Hardy, C. R. A study of the feasibility of accessing information on Prestel by numeric representation of keywords, M.Sc. thesis, Loughborough University of Technology, 1978.

Information Services Co-ordinating Group, New wave is not quite with us yet, Canadian Library Journal, 1981, 38, 5-10.

Inoue, R. The index system of the CAPTAIN system experimental service, In Viewdata and Videotext 1980-81: A Worldwide Report; White Plains, New York: Knowledge Industry Publications, 1980, 113-21.

Jaynes, J. The Origins of Consciousness in the Breakdown of the Bicameral Mind, Boston: Houghton Mifflin, 1976.

Jensen, K. & Ball, A. J. S. A novelty whose time has come, Canadian Library Journal, 1981, 38, 199-205.

King, D. W. & Bryant, E. C. The Evaluation of Information Services and Products, Washington, D.C.: Information Resources Press, 1971.

Lancaster, F. W. Information Retrieval Systems: Characteristics, Testing, and Evaluation, New York: John Wiley & Sons, 1968.

Latremouille, S. & Lee, E. The design of videotex tree indices: the use of descriptors and the enhancement of single index pages, In Telidon Behavioural Research 2, Department of Communications, Ottawa, 1981, 63-112.

Lee, E. The optimum number of alternatives to display on an index page in an interactive Telidon database, In Telidon Behavioural Research 1, Department of Communications, Ottawa, 1980, 128-49.

Lee, E. & Latremouille, S. Evaluation of tree-structured organization of information on Telidon, In Telidon Behavioural Research 1, Department of Communications, Ottawa, 1980, 231-42.

Lochovsky, F. H. & Tsichritzis, D. C. Interactive Query Languages for External Data Bases, Toronto: Computer Systems Research Group, University of Toronto, 1981.

Mahony, S.; Damartino, N. & Stengel, R. Keeping PACE with the New Television: Public Television and Changing Technology, New York: Carnegie Corporation of New York, 1980.

Martin, J. Design of Man-Computer Dialogues, New York: Prentice-Hall, 1973.

Mason, R. O. & Mitroff, I. I. A program for research on management information systems, Management Science, 1973, 19, 475-87.

Mauerhoff, G. Database applications for videotex, Proceedings, Videotex '81 International Conference & Exhibition, 20-22 May 1981; Toronto: Infomart, 1981, 411-16.

- McEwen, S. A. An investigation of user search performance on a Telidon information retrieval system, In Telidon Behavioural Research 2, Department of Communications, Ottawa, 1981, 33-62.
- Meadow, C. T. The Analysis of Information Systems, 2nd ed., Los Angeles: Melville Publishing Company, 1973.
- Miller, G. A. The Psychology of Communication: Seven Essays, New York: Basic Books, 1967.
- Miller, L. A. & Thomas, J. C. Behavioral issues in the use of interactive systems, International Journal of Man-Machine Studies, 1977, 9, 509-36.
- Moran, T. P. An applied psychology of the user, Computing Surveys, 1981, 13, 1-11.
- Muter, P.; Treurniet, W. & Phillips, D. Computer-aided learning and videotex, In Proceedings, Third Canadian Symposium on Instructional Technology, Vancouver, 27 February 1980, 319-25.
- Nickerson, R. S. Why interactive systems are sometimes not used by people who might benefit from them, International Journal of Man-Machine Studies, 1981, 15, 469-83.
- Online Computer Library Center Channel 2000: Description and Findings of a Viewdata Test, Conducted by OCLC in Columbus Ohio, October-December 1980; Dublin, Ohio: OCLC Inc., 1981.
- Parkhill, D. Videotex 1980: state of the art, In Richard Larratt (ed.) Inside Videotex: The future... now, Proceedings of a seminar held 13-14 March 1980, Toronto: Infomart; 1980, 12-18.
- Phillips, D. Telidon, the Canadian videotex system, and its implications for information science, In Telidon Behavioural Research 1, Department of Communications, Ottawa, 1980a, 2-8.
- Phillips, D. The implications of Telidon for information science, In Telidon Behavioural Research 1, Department of Communications, Ottawa, 1980b, 27-39.
- Pollard, D. & James, P. Aspects of indexing on Prestel, In Proceedings of the Ninth International Symposium on Human Factors in Telecommunications, Red Bank, New Jersey, 29 Sep - 3 Oct 1980; 9-15.
- Ramsey, H. R. & Atwood, M. E. Human Factors in Computer Systems: A Review of the Literature, 3 vols., (MS 2039) Washington, D. C.: Journal Supplement Abstract Service, 1980.

- Rimmer, T. Viewdata — interactive television, with particular emphasis on the British Post Office's PRESTEL. Paper presented at the Annual Meeting of the Association for Education in Journalism, Houston, Texas, August 1979.
- Robertson, G. A.; McCracken, D. & Newell A. The ZOG approach to man-machine communication, International Journal of Man-Machine Studies, 1981, 14, 461-88.
- Schabas, A. H.; Williamson, N. J. & Eines, T. T. Keyword Indexing, Department of Communications, Ottawa, 1981.
- Shackel, B. Dialogues and language — can computer ergonomics help?, Ergonomics, 1980, 23, 857-80.
- Shneiderman, B. Software Psychology: Human Factors in Computer and Information Systems, Englewood Cliffs, New Jersey: Prentice-Hall, 1980.
- Stewart, T.; Cox, B. & Partners Ltd. Human factors in videotex, In Proceedings, 4th International Online Information Meeting, London, 9-11 December 1980; New Jersey: Learned Information, 1980, 87-95.
- Thomas, K. & Schabas, A. H. Implications of Keyword Access for Telidon Software, Department of Communications, Ottawa, 1981.
- Tombaugh, J. W. & McEwen, S. A. Comparison of two information retrieval methods on videotex: tree-structure versus alphabetical directory, In Proceedings of Human Factors in Computer Systems, Gaithersburg, Maryland, 15-17 March 1982.
- Tydemann, J. & Zwimpfer, L. Videotex in the United States . . . toward information diversity, In Proceedings, Videotex '81 International Conference & Exhibition, 20-22 May 1981; Toronto: Infomart, 1981, 341-51.
- van Rijsbergen, C. J. Information Retrieval, London: Butterworths, 1975.
- Whalen, T. & Latremouille, S. The effectiveness of a tree-structured index when the existence of information is uncertain, In Telidon Behavioural Research 2, Department of Communications, Ottawa, 1981, 3-12.
- Williams, E. Who's who on the Prestel database, In Rex Winsbury (ed.) Viewdata in Action: A Comparative Study of Prestel, London: McGraw-Hill, 1981, 99-106.
- Williamson, N. J. Viewdata systems: designing a database for effective user access, Canadian Journal of Information Science, 1981, 6, 1-14.

- Williges, R. C. & Williges, B. H. Modeling the human operator in computer-based data entry, Human Factors, 1982, 24, 285-99.
- Wilson, L. Bell Canada's VISTA project, In Richard Larratt (ed.) Inside Videotex: The future . . . now, Proceedings of a seminar held 13-14 March 1980; Toronto: Infomart, 1980, 76-80.
- Winsbury, R. Electronic editing: database and page design, In Rex Winsbury (ed.) Viewdata in Action: A Comparative Study of Prestel, London: McGraw-Hill, 1981, 145-54.
- Witkin, H. A.; Dyk, R. B.; Farnsworth, H. F.; Goodenough, D. R. & Karp, S. A. Psychological Differentiation, Potomac, Maryland: Lawrence Erlbaum, 1974.
- Witkin, H. A.; Goodenough, D. R. & Oltman, P. K. Psychological differentiation: current status, Journal of Personality and Social Psychology, 1979, 37, 1127-45.
- Woolfe, R.; Cox, B. & Partners Ltd. Videotex: online services for the mass market, In Proceedings, 4th International Online Information Meeting, London, 9-11 December 1980; New Jersey: Learned Information, 1980, 57-62.

APPENDIX A

GLOSSARY OF VIDEOTEX TERMS

Videotex - The internationally adopted generic term for hybrid computer-television communications services. Videotex exists in both active and passive forms and can be operated on a variety of transmission media, including radio wave, coaxial cable, copper wire, and optical fibre.

Broadcast videotex - The generic term for one-way, or passive, videotex -- also called teletext. Under this arrangement, a limited number of information pages are broadcast as a package at regular cyclic intervals and the receiver captures them as they roll by for subsequent accessing.

Interactive videotex - The generic term for two-way videotex -- also called viewdata. Interactive videotex allows a user to access selectively a theoretically unlimited number of information pages, and/or other electronic services, at a time of the user's choosing.

Prestel - A first-generation videotex technology and the proprietary name of the interactive videotex service pioneered by the British Post Office. Prestel was introduced to the public in 1976 and as of 1981 contained approximately 160,000 database pages.

Telidon - A second-generation videotex technology developed by the Canadian Department of Communications and first demonstrated in public in 1978. Telidon has been publicized as being system-independent and as being capable of generating high-resolution geometric graphics from a smaller memory store than that required by first-generation, low-resolution, mosaic-graphics systems.

Information provider (IP) - Any private or public electronic publisher who supplies to a common carrier information or services suitable for videotex applications.

Page - The basic quantitative unit of videotex displays, a page is one screenful of information. The maximum capacity of a videotex page for television display is 20 lines of 40 characters each, or 800 characters.

Information pages - The essential ingredients of an information database. Information pages which are closely linked to deal with a unitary concept or body of text are called an information document. Simply described, a page is a physical unit, whereas a document is a logical unit.

Indexing pages - Pages which allow a system user to access information pages. Indexing pages contain information about the location of information pages.

APPENDIX B

EXAMPLE OF A HIERARCHICAL SEARCH

Suppose that the system user were looking for information concerning the quality-point average required for a student to be placed on the Dean's List. After selecting "Start Tree Search" on the root page, PAGE 0000, the following menu of general information categories would be displayed:

0100

INFORMATION CATEGORIES: What general category are you interested in?

- 1 - REGULATIONS Admission procedures, definitions,
general academic regulations, and registration details.
- 2 - FACULTIES AND PROGRAMS General requirements for degree and
diploma programs offered at Saint Mary's University.
- 3 - CONTINUING EDUCATION Credit and non-credit upgrading programs
for part-time students, teachers, professionals, and other adults.
- 4 - FINANCIAL INFORMATION Fees and financial information, including
academic awards, scholarships, bursaries, and loans.
- 5 - ACADEMIC DEPARTMENTS Specific departmental requirements and
descriptions of all courses offered by university departments.
- 6 - CALENDAR OF EVENTS 1982-83 Key dates in the academic year, e.g.
administrative deadlines, examination periods, and holidays.

Surmising that information on the Dean's List is most likely to found under General Academic Regulations, the first information category is selected by keying "1 ENTER". This results in a display of the sub-topics in Regulations:

0110

REGULATIONS: Which regulations topic are you interested in?

- 1 - Admission Pre-requisites and Procedures
- 2 - Definitions
- 3 - Academic Regulations
- 4 - Registration
- 5 - STOP information search

Our searcher keys "3 ENTER" and gets the Index of Academic Regulations:

1120

INDEX OF ACADEMIC REGULATIONS: Which regulation are you interested in?

- | | |
|---------------------------------------|--|
| 1 - Number of Courses in a Year | 16 - Withdrawing from a Course |
| 2 - Auditing Courses | 17 - Retaking a Course |
| 3 - Academic Advising | 18 - Withdrawal for Academic Reasons |
| 4 - Grading System | 19 - Academic Responsibility |
| 5 - Undergraduate Ratings (Grades) | 20 - Advanced Standing |
| 6 - Quality Point Average | 21 - Transfer Credit |
| 7 - Standing Required for Continuance | 22 - Advanced Standing by Examination |
| 8 - Examinations | 23 - Second Undergraduate Degree |
| 9 - Evaluations | 24 - Certificate of Honors Equivalency |
| 10 - Special Examinations | 25 - Convocation Dates, Degrees |
| 11 - Academic Appeals | 26 - Degree or Diploma in Absentia |
| 12 - Credit without Final Examination | 27 - Distinctions |
| 13 - Course Changes | 28 - University Medals |
| 14 - Declaration/Change of Major | 29 - Dean's List |
| 15 - Procedure for Changing Faculty | 30 - Transcripts |
-

Upon keying "29 ENTER", the target page is displayed:

1275

REG #29 DEAN'S LIST

At the end of the academic year, full-time students whose quality point average indicates high academic achievement will have their names placed on the Dean's List by the Dean of the Faculty. To qualify for this recognition, students must have taken at least five full courses (or the equivalent) during the academic year and have achieved a quality point average of 3.50 or higher. Placement on the Dean's List will be recorded on the students' transcript.

FOR LAST REGULATION, PRESS "1 ENTER".
FOR A NEW TOPIC, PRESS "2 ENTER".

From this point, our searcher has the option of directly accessing the next and last information document at this level or calling up PAGE 1398 for a new topic:

1398

NEW TOPIC: Where would you like to go from here?

- 1 - Index of Academic Regulations
 - 2 - NEW TOPIC IN REGULATIONS
 - 3 - NEW INFORMATION CATEGORY
 - 4 - NEW KEYWORD
 - 5 - STOP INFORMATION SEARCH
-

APPENDIX C

EXAMPLE OF A KEYWORD SEARCH

Suppose that a system user were looking for information concerning the quality-point average required for a student to be placed on the Dean's List. After selecting "Start Keyword Search" on the root page, PAGE 0000, the following menu of alphabetic ranges would be displayed:

8000
KEYWORD RANGES: Which alphabetic range contains your target word or phrase?

1 - A	14 - N
2 - B	15 - O
3 - C	16 - P
4 - D	17 - Q
5 - E	18 - R
6 - F	19 - S
7 - G	20 - T
8 - H	21 - U
9 - I	22 - V
10 - J	23 - W
11 - K	- X
12 - L	25 - Y
13 - M	26 - Z

At this point, our searcher must think of a keyword or key-phrase to use as a search term and then select the appropriate alphabetic range. One possibility would be "Academic regulations", but a more specific term such as "Quality-point average" or even "Dean's list" would be much better. Assuming that "Dean's list" is chosen as the search term, the correct selection would be "4 ENTER", which results in a display of the first page of keywords beginning with the letter "D":

8040

KEYWORDS: D...

-
- 1 = Dante Alighieri
 - 2 = Data processing
 - 3 = Data structures
 - 4 = Deadlines
 - 5 = Dean (defined)
 - 6 = Dean's list
 - 7 = December events
 - 8 = Decision making
 - 9 = Declaration of major
 - 10 = Definitions (academic)
 - 11 = Defoe, Daniel
 - 12 = Deformable bodies
 - 13 = Degree requirements
 - 14 = Degree/diploma distinctions
 - 15 = Degree/diploma in absentia
 - 16 = Degrees/diplomas awarded
 - 17 = NEXT PAGE OF KEYWORDS
-

FOR KEYWORDS STARTING:

Den . . . --> 8041
Dre . . . --> 8042

Because there are more than two pages of keywords beginning with the letter "D", a mini-directory of the remaining pages is included, so that if the searcher does not wish to browse through these pages by keying "17 ENTER", the desired page can be directly accessed by keying in its four-digit page number. In this case, the search term is on the first page displayed and the searcher should now key "6 ENTER". Because only one document has been referenced under the phrase "Dean's list", the target information page is immediately displayed:

1275

REG #29 DEAN'S LIST

At the end of the academic year, full-time students whose quality point average indicates high academic achievement will have their names placed on the Dean's List by the Dean of the Faculty. To qualify for this recognition, students must have taken at least five full courses (or the equivalent) during the academic year and have achieved a quality point average of 3.50 or higher. Placement on the Dean's List will be recorded on the students' transcript.

FOR LAST REGULATION, PRESS "1 ENTER".
FOR A NEW TOPIC, PRESS "2 ENTER".

From this point, our searcher has the option of directly accessing the next and last information document at this level or calling up PAGE 1398 for a new topic:

1398

NEW TOPIC: Where would you like to go from here?

- 1 - Index of Academic Regulations
 - 2 - NEW TOPIC IN REGULATIONS
 - 3 - NEW INFORMATION CATEGORY
 - 4 - NEW KEYWORD
 - 5 - STOP INFORMATION SEARCH
-

Because only one document in the database was referenced under the key-phrase "Dean's list", selection of this term on PAGE 8040 resulted in the immediate display of this document. If "Quality-point average" had been used as the search term however, an extra step would have been required. For example, selecting the letter "Q" on PAGE 8000 would call up PAGE 8170:

8170

KEYWORDS: Q...

- 1 = Quality of work life
- 2 = Quality points (course-work)
- 3 = Quality-point average
- 4 = Quantitative methods
- 5 = Quebec applicants

END OF KEYWORDS FOR THIS LETTER

Keying "3 ENTER" now results in a display of the titles of all documents in the database referenced under "Quality-point average", together with the corresponding page numbers of these documents:

KEYWORD: Quality-point average

8602

PAGE: DOCUMENT TITLE:

1150 - REG #6 QUALITY POINT AVERAGE

1265 - REG #27 DISTINCTIONS

1270 - REG #28 UNIVERSITY MEDALS

1275 - REG #29 DEAN'S LIST

END OF PAGES FOR THIS KEYWORD

By keying in the four-digit page number, 1275, the target document would now be displayed.

APPENDIX D

PERSONAL DATA QUESTIONNAIRE

Please print.

Name _____

Sex _____

Left or right handed _____

Age _____

Academic major _____

A. Have you had any previous computer experience? Yes / No
If "Yes", please describe briefly:

B. Have you ever used a videotex system (Questel, Telidon)? . . Yes / No
If "Yes", please indicate amount of experience in hours:

C. Have you used the University's Academic Calendar before? . . Yes / No

APPENDIX E

VIDEOTEX ATTITUDE QUESTIONNAIRE

For questions 1 to 4, you are to rate the potential value of computerized video systems as media for presenting and accessing information services of various kinds. Circle the number of the response which best describes your opinion of the stated proposal.

1. Suppose that the University were to make available to staff and students a computerized calendar-information service. What would be your reaction to such a proposal?
 - 1 - unfavourable
 - 2 - unsure or no opinion
 - 3 - favourable

2. Suppose that the University were to make available to staff and students a computerized calendar service which would also permit registration, grade reporting, and other academic administration by means of an interactive terminal? What would be your reaction to such a proposal?
 - 1 - unfavourable
 - 2 - unsure or no opinion
 - 3 - favourable

3. Suppose that the University were to make a service such as that described in Question 2 available to the general public through home televisions equipped with keypads or other input devices. What would be your reaction to such a proposal?
 - 1 - unfavourable
 - 2 - unsure or no opinion
 - 3 - favourable

4. Suppose that a number of private and public organizations were to make available to the general public, through home televisions equipped with keypads or other input devices, a wide variety of interactive, computer services. What would be your reaction to such a proposal?
 - 1 - unfavourable
 - 2 - unsure or no opinion
 - 3 - favourable

APPENDIX F

SEARCH-METHOD RATING SCALE

For questions 1 to 10, you are to rate the search procedure that you used in this experiment -- that is, either the TREE or KEYWORD procedure -- on the performance features described. Indicate your choice of rating on each feature by circling the number of the applicable statement. When considering your choice, try not to be influenced by any opinions you may have formed concerning the content or possible applications of the information system, namely, calendar information. You are being asked to evaluate only the search procedure.

1. In comparison to looking up information in the printed calendar, do you find this method more, or less, interesting?
 - 1 - much less interesting
 - 2 - less interesting
 - 3 - slightly less interesting
 - 4 - no difference
 - 5 - slightly more interesting
 - 6 - more interesting
 - 7 - much more interesting

2. In comparison to looking up information in the printed calendar, do you find this search method more, or less, efficient?
 - 1 - much less efficient
 - 2 - less efficient
 - 3 - slightly less efficient
 - 4 - no difference
 - 5 - slightly more efficient
 - 6 - more efficient
 - 7 - much more efficient

3. How would you rate this search procedure on ease of learning?

- 1 - very difficult to learn
- 2 - difficult to learn
- 3 - somewhat difficult to learn
- 4 - in-between
- 5 - somewhat easy to learn
- 6 - easy to learn
- 7 - very easy to learn

4. How would you rate this search procedure on overall ease of use?

- 1 - very difficult to use
- 2 - difficult to use
- 3 - somewhat difficult to use
- 4 - in-between
- 5 - somewhat easy to use
- 6 - easy to use
- 7 - very easy to use

5. How would you rate this search procedure on speed of finding information?

- 1 - very slow
- 2 - slow
- 3 - somewhat slow
- 4 - in-between
- 5 - somewhat fast
- 6 - fast
- 7 - very fast

6. How efficient is this search procedure?

- 1 - always takes a lot of pages to find information
- 2 - usually takes a lot of pages to find information
- 3 - tends to take a lot of pages to find information
- 4 - in-between
- 5 - tends to take only a few pages to find information
- 6 - usually takes only a few pages to find information
- 7 - always takes only a few pages to find information

7. How would you rate this search procedure on its thoroughness in finding target information?

- 1 - fails to find any target information
- 2 - fails to find most target information
- 3 - fails to find a lot of target of information
- 4 - in-between
- 5 - succeeds in finding a lot of target information
- 6 - succeeds in finding most target information
- 7 - succeeds in finding all target information

8. How would you rate this search procedure on error-proneness?

- 1 - very easy to make search errors
- 2 - easy to make search errors
- 3 - somewhat easy to make search errors
- 4 - in-between
- 5 - somewhat difficult to make search errors
- 6 - difficult to make search errors
- 7 - very difficult to make search errors

9. How would you rate this search procedure with respect to its ability to solve different kinds search problems?

- 1 - very restricted in the kinds of problems it can solve
- 2 - restricted in the kinds of problems it can solve
- 3 - somewhat restricted in the kinds of problems it can solve
- 4 - in-between
- 5 - somewhat unrestricted in the kinds of problems it can solve
- 6 - unrestricted in the kinds of problems it can solve
- 7 - totally unrestricted in the kinds of problems it can solve

10. To what extent does this search procedure allow you a choice of search routes to target information?

- 1 - search procedure completely controls search route
- 2 - search procedure controls search route
- 3 - search procedure somewhat controls search route
- 4 - in-between
- 5 - user somewhat controls search route
- 6 - user controls search route
- 7 - user completely controls search route

* * * * *

APPENDIX G

EXPERIMENTER'S INSTRUCTIONS TO SUBJECTS

GENERAL INSTRUCTIONS

The purpose of this experiment is to evaluate two different procedures for finding information in a computerized information system. The experiment will take approximately 1 1/2 hours to complete. To ensure you understand the conditions governing your participation, I would like to remind you:

- (a) that you are free to discontinue your participation in the experiment at any time,
- (b) that individual data will be treated as confidential, and
- (c) that if you are dissatisfied in any way with experimental procedures you should discuss your concerns with the Departmental Chairperson.

REQUIREMENTS OF THE EXPERIMENT

There are four phases: (1) you'll be asked to answer a few questions about computerized information systems and write two very short aptitude tests; (2) you'll receive instruction in how to use the calendar information system and one of the search procedures; (3) then, you'll be given a series of test problems to solve on your own; and (4) at the end of the test problems, I'll ask you to rate the search procedure on a number of performance dimensions.

Do you have any questions at this point concerning the general framework of the experiment?

(ADMINISTER PRETEST QUESTIONNAIRE AND ABILITY TESTS.)

THE CALENDAR INFORMATION SYSTEM

As you can see from the display, the information system to be used in this experiment is based on the Academic Calendar. The contents and organization are generally similar to the contents and organization of the print version except that about 60% of the academic departments are not represented and information is displayed on electronic pages rather than paper pages.

A page in this system is one screenful of text and shows (1) in the

top right corner, a PAGE NUMBER (this is PAGE 0000) . . . (2) then a body of text . . . and (3) a MENU or a KEYING INSTRUCTION for advancing to other pages. There are two ways of retrieving information pages from computer memory in this system. One is by carrying out a TREE SEARCH; the other is by a KEYWORD SEARCH. Very shortly, you will learn how to do one of them. But first you'll have to know which buttons to push.

All input is numerical and so the keypad you'll be using is based on 10 number keys (the digits 0 to 9), an oversized ENTER key, and four function keys -- DELETE, HELP, HIT, and BACK-UP (the key with the minus sign).

(BRANCH HERE ACCORDING TO SEARCH PROCEDURE TO BE USED.)

INSTRUCTIONS FOR HIERARCHICAL SEARCH PROCEDURE

1. Calling up pages: There are several ways to call up pages: one is by keying in a number offered in a MENU or a KEYING INSTRUCTION, and another is by keying in the four-digit PAGE NUMBER of the page you want -- if you know what the page number is. In any case, every keypad entry must always be followed by the ENTER keystroke in order to be executed.

You'll be using the TREE search procedure . . . so key the appropriate command to start. . . . Here we have a new MENU of numbered INFORMATION CATEGORIES from which you may select one. . . . Let's try item 5 in this menu, ACADEMIC DEPARTMENTS At this point, let's suppose you wanted information about some topic in the Biology Department . . . what would you key in? . . . And now, if you wanted information about undergraduate courses? . . . Which level of courses? . . . And finally, which course? . . . As just demonstrated then, I hope you understand that to reach an INFORMATION PAGE you must advance through a series of MENU PAGES.

If you had known the page number beforehand, you could have called up this page from PAGE 0000 simply by keying in the four-digit page number. Chances are though that you'll rarely know the page numbers of individual information documents in advance.

2. Backing-up: You now know how to go forward in this system. You can also back up to pages you've already looked at. Let's say you want to see the menu of courses for this level again and pick another course. . . . Key "- ENTER". . . . Notice that "- ENTER" has the same effect as "-1 ENTER"; but the general form of the BACK-UP command, if you want to back up N pages, is "-N ENTER" . . . for example, "-3 ENTER" to back up 3 pages, "-10 ENTER" to back up 10, and so on.

3. What to do if you make a mistake:

Key in any number but don't press ENTER If you just want to change an entry and you haven't yet pressed the ENTER key, all you have to do is use the DELETE key.

Now, key in any number that is not offered as a selection and press ENTER. In this case, you get an error message.

Finally, if you become hopelessly muddled or lost, you can use the special command "0 ENTER", at any time, to restart at PAGE 0000 . . . try it . . . And, as you see, we're back where we started.

4. Getting HELP: What if you forget what to do? . . . Try HELP . . . (PAGE 9998 displayed). Here you see a summary of system-operating features that can be called up at any time from any page. Any questions?

5. Organization and search strategy: It may help you in carrying out a TREE search if you think of the information documents and pages as a

collection of articles in a large book. The tree system is like a hierarchy of Tables of Contents. At the top of the hierarchy, you will find a menu of the most general categories, similar to a general Table of Contents; at the next lower level, you find more specialized menus, or Tables of Contents, which lead eventually to particular "articles" or information pages. Retrieving information is thus a matter of moving down through a series of increasingly more specialized Tables of Contents until you find the document and page you are looking for.

But, because menu pages are arranged in the order of "most general" categories to "most specific", the best strategy when you begin a search is to select a general search term for your target information. In other words, ask yourself, "In what general category am I likely to find this information?". It is very important to select the right information category at the beginning of your search; otherwise you'll be wasting time and effort looking down a blind alley. So remember, a general search term to start!

6. Demonstration: Let's assume that you'd like to find a course description for an introductory course in logic. We start our search by selecting the appropriate search procedure. Here is our menu of INFORMATION CATEGORIES; which one do we want? . . . Remember -- which general category seems most likely to contain this information? . . . And now which menu item?

(CONTINUE, WITH GUIDANCE, UNTIL PROBLEM SOLVED.)

7. Browsing: If, at this point, you wanted to browse through the 200-level courses in Philosophy, you would follow the FOR NEXT COURSE, keying instruction at the bottom of the page. Continuing in this fashion, you could browse through all the information pages at this level and eventually return to the page you started on.

8. New topics: If, however, you wished to break out of this browsing loop and obtain information on a new topic, you would follow the NEW TOPIC keying instructions. . . . Try "2 ENTER" . . . You see that from here you have several options.

9. Experimental conventions: These are the basics of how to go about a TREE search. . . . Do you have any questions? . . . Let's return to the page on Basic Logic for a moment. There are two conventions which are extremely important for gathering experimental data.

First of all, whenever you find and identify an information page you are looking for -- and you must decide whether or not it's a target page -- you must indicate you have found it by keying "HIT ENTER" on that page. The sign in front of you is there as a reminder. Try it. . . . The system will reply as shown and you can then make another entry. It is very important that you record a HIT while you are on the target page, and not later on some other page.

The second convention is this. As soon as you have finished a search, either because you have found all the target pages you can or because

you have given up, you must end the search by keying "98 ENTER". Again, the sign in front of you is there as a reminder. Try it . . . As you see, the problem you were working on is closed off and a new search is begun with this prompt to enter the number of the next problem. The system will then start you off on PAGE 0000 again.

Thus the typical sequence, if you are required to find only one page of information, will be (1) enter problem number, (2) conduct search, (3) if you find the target page, record a HIT and end the search immediately with "98 ENTER", OR, if you don't find the target page, end the search with "98 ENTER" when you have given up. If you are required to find more than one page of information, the procedure is basically the same except that you would record a HIT on every target page you find and would key "98 ENTER" only after finding the last target page.

How will you know if you are supposed to find more than one page? If you are required to look for one page, the problem statement will begin "Find the page." If you are required to look for more than one page, the problem statement will begin "Find as many pages as you can."

Do you understand the conventions for recording an information HIT and for ending a search?

10. Practice problem: I now have a practice problem for you to try on your own. In solving this problem, I'd like you to use the experimental conventions I just explained, that is, (1) enter the problem number when prompted, (2) when you find the answer, key "HIT ENTER" on the target page, and (3) end the search with "98 ENTER". Here is the problem, and remember, HELP is always available to you . . .

(SUBJECT ATTEMPTS PRACTICE PROBLEM, ASSISTED ONLY IF REQUIRED.)

11. Test problems: Test problems will be presented to you on index cards. Please attempt them in order and do not skip any problems. If you accidentally end a search with "98 ENTER", simply enter the same problem number immediately when prompted and continue; otherwise, you may not make a second attempt on a problem once you have started a new one. Some problems will require you to find more than one page and to record more than one HIT. There are also a few problems for which the required information cannot be found. You must decide when a problem is solved or when a search is not worth continuing. If you have made what you consider to be a reasonable effort to find the information but have not been successful, do not hesitate to end the search with "98 ENTER". There are no time limits, but please work as quickly as you can. As a final word of advice, remember that your best strategy in a TREE search is to start with a broad search term. Do you have any questions?

INSTRUCTIONS FOR KEYWORD SEARCH PROCEDURE

1. Calling up pages: There are several ways to call up pages: one is by keying in a number offered in a MENU or a KEYING INSTRUCTION, and another is by keying in the four-digit PAGE NUMBER of the page you want — if you know what the page number is. In any case, every keypad entry must always be followed by the ENTER keystroke in order to be executed.

You'll be using the KEYWORD search procedure . . . so key the appropriate command to start. . . Here we have a new MENU of ALPHABETIC RANGES from which you would select one, the letter that your search term begins with. Let's suppose you wanted information about courses in "Botany". . . What would you key in? . . . Let's try item 2, the letter "B". . . And now, from the list of document titles referenced under "Botany", which document would you choose? . . . As just demonstrated then, I hope you understand that to reach an INFORMATION PAGE you must advance through a series of MENU PAGES.

Once you have a menu of document titles and their page numbers, it's fairly easy to call up a document by keying in the four-digit page number. And this will be the usual way of getting to information pages in a KEYWORD search.

2. Backing-up: You now know how to go forward in this system. You can also back up to pages you've already looked at. Let's say you want to see the document titles listed under "Botany" again. . . Key "- ENTER". . . Notice that "- ENTER" has the same effect as "-1 ENTER"; but the general form of the BACK-UP command, if you want to back up N pages, is "-N ENTER". . . for example, "-3 ENTER" to back up 3 pages, "-10 ENTER" to back up 10, and so on.

3. What to do if you make a mistake:

Key in any number but don't press ENTER. . . If you just want to change an entry and you haven't yet pressed the ENTER key, all you have to do is use the DELETE key.

Now, key in any number that is not offered as a selection and press ENTER. In this case, you get an error message.

Finally, if you become hopelessly muddled or lost, you can use the special command "0 ENTER", at any time, to restart at PAGE 0000 . . . try it . . . and, as you see, we're back where we started.

4. Getting HELP: What if you forget what to do? . . . Try HELP . . . (PAGE 9998 displayed). Here you see a summary of system-operating features that can be called up at any time from any page. Any questions?

5. Organization and search strategy: It might help you in carrying out a KEYWORD search if you think of the information documents and pages as a collection of articles in a large book. The keyword system is like a

back-of-the-book index. You simply look up the term of interest to you in the alphabetical index, find out which "articles" deal with that topic, and turn to the indicated page. Documents will be listed under more than one keyword, and so if your first search term doesn't work you should try a different one.

But, because information documents can be indexed under specific terms as well as general terms, the best strategy when you begin a search is to select a specific keyword for your search term. A general term might get you to the same document, but specific terms will usually be faster. So remember, a specific keyword such as a name, title, or a fairly narrow topic to start!

6. Demonstration: Let's assume that you'd like to find a course description for an introductory course in logic. We start our search by selecting the appropriate search procedure.

Here is our menu of ALPHABETIC RANGES; which letter do we want? . . .
Remember -- a specific keyword . . . And now which menu item?

(CONTINUE, WITH GUIDANCE, UNTIL PROBLEM SOLVED.)

7. Browsing: If, at this point, you wanted to browse through the 200-level courses in Philosophy, you would follow the FOR NEXT COURSE keying instruction at the bottom of the page. Continuing in this fashion, you could browse through all the information pages at this level and eventually return to the page you started on. Notice, however, that the next course is the next 200-level course and not the next logic course. This is because courses are grouped by level within Departments and not by any particular keyword.

8. New topics: If, however, you wished to break out of this browsing loop and obtain information on a new topic, you would follow the NEW TOPIC keying instructions. . . . Try "2 ENTER" . . . You see that from here you have several options. At this point, you would probably choose NEW KEYWORD.

9. Experimental conventions: These are the basics of how to go about a KEYWORD search. Do you have any questions? . . . Let's return to the page on Basic Logic for a moment. There are two conventions which are extremely important for gathering experimental data.

First of all, whenever you find and identify an information page you are looking for -- and you must decide whether or not it's a target page -- you must indicate you have found it by keying "HIT ENTER" on that page. The sign in front of you is there as a reminder. Try it . . . The system will reply as shown and you can then make another entry. It is very important that you record a HIT while you are on the target page, and not later on some other page.

The second convention is this. As soon as you have finished a search, either because you have found all the target pages you can or because you have given up, you must indicate the end of the search by keying "98

ENTER". Again, the sign in front of you is there as a reminder. Try it . . . As you see, the problem you were working on is closed off and a new search is begun with this prompt to enter the number of the next problem. The system will then start you off on PAGE 0000 again.

Thus the typical sequence, if you are required to find only one page of information, will be (1) enter problem number, (2) conduct search, (3) if you find the target page, record a HIT and end the search immediately with "98 ENTER", OR, if you don't find the target page, end the search with "98 ENTER" when you have given up. If you are required to find more than one page of information, the procedure is basically the same except that you would record a HIT on every target page you find and would key "98 ENTER" only after finding the last target page.

How will you know if you are supposed to find more than one page? If you are required to look for one page, the problem statement will begin "Find the page." If you are required to look for more than one page, the problem statement will begin "Find as many pages as you can."

Do you understand the conventions for recording an information HIT and for ending a search?

10. Practice problem: I now have a practice problem for you to try on your own. In solving this problem, I'd like you to use the experimental conventions I just explained, that is, (1) enter the problem number when prompted, (2) when you find the answer, key "HIT ENTER" on the target page, and (3) end the search with "98 ENTER". Here is the problem, and remember, HELP is always available to you . . .

(SUBJECT ATTEMPTS PRACTICE PROBLEM, ASSISTED AS REQUIRED.)

11. Test problems: Test problems will be presented to you on index cards. Please attempt them in order and do not skip any problems. If you accidentally end a search with "98 ENTER", simply enter the same problem number immediately when prompted and continue; otherwise, you may not make a second attempt on a problem once you have started a new one. Some problems will require you to find more than one page and "to record more than one HIT". There are also a few problems for which the required information cannot be found. You must decide when a problem is solved or when a search is not worth continuing. If you have made what you consider to be a reasonable effort to find the information but have not been successful, do not hesitate to end the search with "98 ENTER". There are no time limits, but please work as quickly as you can. As a final word of advice, remember that your best strategy in a KEYWORD search is to start with a narrow search term. Do you have any questions?

APPENDIX H

STATISTICAL TABLES

TABLE 11

Multivariate analysis of effects
in solvable and unsolvable problems

Effect	Wilks' Λ	df	Approximate F
Search Method	.58100	(4,37)	6.67*
Search Task	.03364	(8,33)	118.48*
Method x Task	.40533	(8,33)	6.05*

* $p < .001$

TABLE 12

Univariate analysis of Search-Method effects
in solvable and unsolvable problems

Variable	Hypoth MS	Error MS	df	F
Hits	3.592	.201	(1,40)	17.88*
Pages	1.571	60.669	(1,40)	.03
Search time	16032.490	6642.735	(1,40)	2.41
Time per page	31.797	14.903	(1,40)	2.13

* $p < .001$

TABLE 13

Univariate analysis of Search-Task effects
in solvable and unsolvable problems

Variable	Hypoth MS	Error MS	df	F	Adj df ^a	F
Hits	26.736	.120	(1,40)	222.85**	(1,40)	222.85**
Pages	1531.930	.33.193	(2,80)	46.15**	(1,40)	46.15**
Search time	114727.533	2644.012	(2,80)	43.39**	(1,40)	43.39**
Time per page	19.752	2.012	(2,80)	9.81**	(1,40)	9.81*

^a Greenhouse-Geisser adjustment to degrees of freedom to correct for heterogeneity of within-cells variance-covariance matrices.

* $p < .005$

** $p < .001$

TABLE 14

Univariate analysis of Search-Method x Task interactions
in solvable and unsolvable problems

Variable	Hypoth MS	Error MS	df	F	Adj df ^a	F
Hits	1640	.120	(1,40)	5.33 *	(1,40)	5.33 *
Pages	620.651	33.193	(2,80)	18.70 **	(1,40)	18.70 **
Search time	34712.619	2644.012	(2,80)	13.13 **	(1,40)	13.13 **
Time per page	5.440	2.012	(2,80)	2.70	(1,40)	2.70

^a Greenhouse-Geisser adjustment to degrees of freedom to correct for heterogeneity of within-cells variance-covariance matrices.

* $p < .05$

** $p < .001$

TABLE 15a

Two-tailed tests of significance
for comparisons of performance means

Factor	Contrast			Standard error	df	t
	Unipart	Multipart	Unsolvable			
<u>Search Task</u>						
Hits	.76	<	1.89	.079	40	14.30**
Pages	11.26	<	18.35	.821	80	8.64**
	11.26		<	23.28	80	6.89**
		18.35	<	23.28	80	2.82*
Search time	114.74	<	200.66	8.56	80	10.04**
	114.74		<	209.26	80	7.38**
		200.66	<	209.26	80	.54
Time/page	10.22	<	10.98	.287	80	2.65**
	10.22		<	9.61	80	-1.81
		10.98	<	9.61	80	-4.26**

* $p < .05$

** $p < .01$

TABLE 15b

Two-tailed tests of significance
for comparisons of performance means

Factor	Contrast		Standard error	df	t
	Keyword	Hierarchy			
<u>Method x Task</u>					
Unipartite tasks					
Hits	.88	> .64	.107	40	2.24
Pages	8.95	< 13.57	1.78	80	2.60*
Search time	89.35	< 140.13	15.87	80	3.20**
Multipartite tasks					
Hits	2.18	< 1.59	.107	40	5.52**
Pages	16.45	< 20.24	1.78	80	2.13
Search time	170.42	< 230.91	15.87	80	3.81**
Unsolvable tasks					
Pages	27.82	< 18.73	1.78	80	-5.11**
Search time	231.05	< 187.46	15.87	80	-2.75*

* $P < .05$

** $P < .01$

TABLE 16

Multivariate analysis of effects
in solvable problems

Effect	Wilks' Λ	<u>df</u>	Approximate <u>F</u>
Search Method	.47204	(4,37)	10.35*
Search Task	.12209	(4,37)	66.51*
Method x Task	.80738	(4,37)	2.21

* $p < .001$

TABLE 17
Univariate analysis of Search-Method effects
in solvable problems

Variable	Hypoth MS	Error MS	df	F
Hits	3.592	.201	(1,40)	17.88**
Pages	372.038	26.160	(1,40)	14.22*
Search time	65007.550	4936.192	(1,40)	13.17*
Time per page	9.282	11.182	(1,40)	.83

* $P < .01$

** $P < .001$

TABLE 18

Univariate analysis of Search-Task effects
in solvable problems

Variable	Hypoth MS	Error MS	df	F
Hits	26.736	.120	(1,40)	222.85**
Pages	1053.646	14.430	(1,40)	73.02**
Search time	155040.922	1563.094	(1,40)	99.19**
Time per page	12.261	1.684	(1,40)	7.28*

* $p < .05$

** $p < .001$

TABLE 19

Analysis of variance for regression
of retrieval successes on spatial memory scores

Source	SS	df	MS	F
Regression	1.036	1	1.036	8.68*
Residual	4.778	40	.119	
Total	5.814	41		

* $p < .01$

TABLE 20

Analysis of variance for regression
of pages accessed on ideational fluency scores

Source	SS	df	MS	F
Regression	62.49	1	62.49	4.35*
Residual	574.21	40	14.35	
Total	636.71	41		

* $p < .05$

TABLE 21

Analysis of variance for regression
of retrieval successes on spatial memory scores
(Hierarchical search)

Source	SS	df	MS	F
Regression	.7716	1	.7716	9.78*
Residual	1.4994	19	.0789	
Total	2.2710	20		

* $p < .01$

TABLE 22

Analysis of variance for regression
of pages accessed on spatial memory scores
(Hierarchical search)

Source	SS	<u>df</u>	MS	<u>F</u>
Regression	65.49	1	65.49	4.67*
Residual	266.93	19	14.04	
Total	332.42	20		

* $p < .05$

TABLE 23

Analysis of variance for regression
of pages accessed on ideational fluency scores
(Keyword search)

Source	SS	df	MS	F
Regression	82.39	1	82.39	7.96*
Residual	196.67	19	10.35	
Total	279.07	20		

* $p < .05$

TABLE 24

Multivariate analysis of differences in ratings
between search methods

Effect	Hotelling's T^2	df	Approximate F
Search Method	.24210	(10,31)	.75

TABLE 25

Non-parametric tests of association
between subject characteristics
and pre-exposure attitudes toward videotex

Variables	df	χ^2	Adj χ^2	ϕ Coefficient
<u>Gender</u>				
Computerized calendar	1	.20	.01	.07
Calendar + transactions	1	.01	0	.02
Videotex calendar	1	.47	.12	.12
Commercial videotex	1	.49	.09	.11
<u>Computer Experience</u>				
Computerized calendar	1	.02	0	.02
Calendar + transactions	1	4.89	3.30	.34
Videotex calendar	1	.05	0	.03
Commercial videotex	1	.02	0	.02