since its foundation in 1945 has established a reputation as a leading journal in the field. It publishes original articles on most aspects of documentation, librarianship and information science, both theoretical and practical. It is concerned with both the processes of information handling and the organisation thereof, i.e. with the extraction, analysis, characterisation, classification, storage, retrieval and dissemination of information recorded in any format; with the institutions and systems, general or special, which handle information; with the social environment of information processing and communication; and with the nature of information itself. The Journal gives special attention to articles treating or illustrating general principles, and presenting new information management techniques. Papers based on research or practice are both accepted, as are those with a historical slant.

The *Journal of Documentation* is published quarterly in March, June, September and December. Annual Subscription: UK £65 (£45 members); Overseas £75 (£50 members). Orders should be addressed to Aslib, The Association for Information Management, 20–24 Old Street, London EC1V 9AD. No bookseller's or other discount can be allowed on members' rates.

The Journal is published by Aslib. It is edited by an Editorial Board consisting of:

All editorial communications should be addressed to the Managing Editor, R. T. Kimber, The Queen's University of Belfast, Belfast, Northern Ireland BT9 5EQ.
For 'Notes for contributors', please see inside back cover.

THE
Journal of Documentation

VOLUME 45 NUMBER 3 SEPTEMBER 1989

A BEHAVIOURAL APPROACH TO INFORMATION RETRIEVAL SYSTEM DESIGN

DAVID ELLIS

Department of Information Studies, University of Sheffield, Western Bank, Sheffield, S10 2TN

A behavioural approach to information retrieval system design is outlined based on the derivation of a behavioural model of the information seeking patterns of academic social scientists. The information seeking patterns of a variety of academic social scientists were broken down into six characteristics: starting, chaining, browsing, differentiating, monitoring, and extracting. These characteristics constitute the principal generic features of the different individual patterns, and together provide a flexible behavioural model for information retrieval system design. The extent to which these characteristics are available on existing systems is considered, and the requirements for implementing the features on an experimental system are set out.

1. INTRODUCTION

ALTHOUGH THE ORIGINS of modern information retrieval research can be traced back to 1953 it is the series of tests carried out at, or in association with, the Cranfield Institute of Technology from 1957 which represent the real beginnings of information retrieval research as a sub-discipline of library or information studies [1]. The Cranfield tests provided the theoretical framework within which the sub-discipline developed and the procedures adopted in those tests have informed a whole tradition of research into the design and testing of information retrieval systems [2–5]. The emphasis in this tradition has been on controlled laboratory tests of indexing devices and systems, with accompanying restrictive assumptions concerning the cognitive and behavioural features of the environment within which the information retrieval system operates [2, 3, 6].

It is only comparatively recently that serious interest has been shown in alternative approaches to information retrieval system design, and most studies which have adopted an alternative approach have focussed on the cognitive aspects of the retrieval interaction [7]. Examples include information retrieval through man-machine dialogue [8–10], anomalous-states-of-knowledge based retrieval [11–16], stereotype based fiction retrieval [17–19], automating the intermediary function [20–23], and utilising domain knowledge [24–26].

An alternative approach to the problem of information retrieval system design is to focus on behavioural aspects of the retrieval interaction, specifically, on how researchers interact with their information sources [27, 28]. If researchers' information seeking patterns are broken down into their basic behavioural characteristics – and the retrieval system is provided with facilities that reflect those characteristics – then users should be able to recreate their own information seeking patterns while interacting with the system. This paper describes the results of a study which adopts such an approach to questions concerning information retrieval system design in relation to the information seeking activities of academic social scientists [27].

2. THE DERIVATION OF THE BEHAVIOURAL MODEL

There are numerous studies of the characteristics of social science information and the information seeking activities of social scientists [29–43] and this literature has been extensively reviewed [44–51]. However, despite representing useful background, none of these studies provide detailed accounts of the perceptions of academic social scientists of their own information seeking activities, from their point of view, and as a whole. It was this type of micro-level information about the activities and perceptions of the academic social scientists which was considered necessary for a detailed analysis of information retrieval system requirements to be possible. Therefore, such accounts needed to be obtained first hand from interviews with active researchers.

For convenience, it was decided to interview social scientists at the University of Sheffield, the sample being identified from:

- MRC/SSRC SAPU and Psychology staff list;
- MRC/SSRC SAPU research assistants;
- Social Science Faculty staff list;
- Library computer search records.

The choice of these groups was based on Glaser and Strauss's [52] technique of theoretical sampling, in that both the choice of initial and subsequent groups to interview was decided with reference to the perceived theoretical and comparative requirements of the project as it progressed.

The choice of the first group of social scientists to interview was influenced by a special characteristic of the University of Sheffield. This was the existence of a special research unit – the MRC/SSRC (now ESRC/MRC) Social and Applied Psychology Unit – within the University. This offered the opportunity

of a potential comparison between the activities of those working in a research unit and those working in a teaching department – the Department of Psychology. All the non-professorial academic staff in the Department of Psychology and all the established researchers in the MRC/SSRC Social and Applied Psychology Unit were contacted and asked if they would be willing to co-operate in this study. Of the twenty-one contacted sixteen were interviewed; of the others two refused to take part, one was on study leave in America, and two were too busy for a convenient interview to be arranged.

It became clear, in the course of these interviews, that the initial choice of sample for interview at the research unit – which had been made using the published list of staff members – had not included other less senior contract research staff of the unit. Therefore, for completeness, four of these researchers were contacted and interviewed. The length of the interviews in this first series ranged from forty-five minutes to two hours with the majority lasting about an hour.

At the outset it had been considered a possibility that, following interviews with the psychologists, interviews with members of a department on what might be conventionally understood as the opposite end of the social science perspective might be attempted – but it was clear from examination of the results of the interviews with the psychologists in both the teaching department and the research unit that there was immense variation in substantive topic and research style within the group. To choose another department on the basis that it would represent a contrast to the 'hardness' of the psychologists seemed unjustifiable. Instead, two members of every department in the faculties of Social Science and Education were contacted – the names taken at random from the University staff list – which consisted of Sociological Studies (two from Sociology and two from Social Administration), Economic and Social History, Economic Studies, Geography, Political Theory and Institutions, Education, and Continuing Education. Due to practical problems one member of Economic Studies and one member of Continuing Education were not interviewed; this was compensated for by the representation of these departments in the next sample. The interviews took the same form as the previous ones with the psychologists and were on the whole a little shorter – lasting between forty-five minutes and an hour. The overall impression again was of variation within as well as between subject groups – and with many similarities across institutional subject boundaries.

The next sample was chosen to investigate the possibility that those social scientists who had had computer based literature searches might differ consistently in their information seeking patterns from those who had not. Although some of the social scientists in the previous samples had had computer based literature searches carried out for them – and had been asked about their experience with these – no systematic attempt had been made to distinguish those who had, and those who had not, had experience of searches of this kind. With the co-operation of the University Library's Social Science Librarian the library's records of online searches carried out in the last several years were examined and the names of the individuals, departments and

search topics noted. All the individuals who had had such searches and were still available within the University were contacted and interviewed. To ensure consistency exactly the same procedure was employed in these interviews as with the other groups. Questions concerning the computer based search were asked at the end of each interview, unless they had come up earlier in the course of the interview. In the event, the only significant difference there seemed to be between this sample and the previous ones was the fact that they had had a computer based search carried out, with varying degrees of success and for varying reasons, and the others generally had not.

In carrying out the interviews the intention was not to set up a situation where the researchers ended up talking rather artificially about their 'information behaviour', or to talk, in isolation, about their use of the library, but rather for them to describe their work and the sorts of activities they engaged in which might be understood as having an information component. Therefore, the questionnaire (Appendix 2) was used as a guide to ensure that no major part of the social scientists' information seeking activities was missed [53], but was not used to 'over determine' the course which an interview might take.

Relevant portions of the taped interviews were transcribed in full, representing something like 2,000 words of transcript per interview. A tape recorder fault led to one interview not being taped; however, from the forty-seven interviews transcribed, the data for analysis comprised circa 250 pages of transcript, or approximately 100,000 words. The analysis of the transcripts took place in part at the same time as the interviewing, in order to provide feedback to help focus future interviews, and in part following completion of the majority of the interviews.

The analysis of the transcripts was closely modelled on that outlined by Glaser and Strauss [52], with further practical guidance on carrying out the analysis being derived from Glaser [54] and Turner [55]. Candidate categories were identified with their appropriate properties. Initially this was done employing cards for significant terms and then listing related terms together; later the transcripts themselves were marked up with previously identified categories and analysed for new candidate categories and properties. As the analysis progressed it became clear that certain categories were more or less synonymous with others and that some candidate categories were better treated as properties of broader categories. In the end six major categories seemed to subsume satisfactorily the important characteristics of the information seeking patterns: starting, chaining, browsing, differentiating, monitoring and extracting.

The derivation of the categories and properties of the categories inductively during the course of data collection and analysis is similar to the procedure Glaser and Strauss [52] outline for generating grounded substantive theory. The construction of the model from a number of semi-independent components corresponds to Diesing's [56] description (derived from Kaplan [57]) of the concatenated structure of holistic theories. The model could also be described as representing an empirically derived [56] and analyst-constructed

[53] typology of the principal characteristics of the social scientists' information seeking patterns.

It was not considered feasible to carry out triangulation of methods, the information seeking activities of academics were too diffuse to make observation practicable, and the likelihood of researchers being either willing or able to record their information seeking activities in the form of diaries seemed remote. However, the interviews and transcripts did seem to provide enough information for a detailed and accurate account of the perceptions of the researchers of their information seeking activities to be possible, and to enable an authentic picture to be constructed of those activities. This seems to be confirmed by the comparison of the results with the findings of other studies of social science information use.

For clarity of exposition the responses of the individual social scientists interviewed are treated together in broad subject groupings:

- MRC/SSRC Social and Applied Psychology Unit
- Psychology
- Education and Continuing Education
- Economics, Economic and Social History, Geography, Politics, Sociology, and Prehistory and Archaeology.

Individual responses are then identified numerically within these groups; in addition, the individual responses are coded alphabetically to enable identification of the sample set from which the individual response was taken. Brief outlines of the subject interests of the social scientists interviewed are given in Appendix 1.

To illustrate the nature of the analysis by which the information seeking characteristics were derived from the information seeking patterns, a few examples of information seeking patterns, or aspects of such patterns, will be analysed into their characteristics. The first example is of a psychologist in a research unit. In this case, it should be noted that the researcher knew little about the topic, nor was there anyone he might have easily contacted who did.

> Three years ago I was working on the area of discourse analysis. I knew that the area of discourse analysis existed but I didn't know anyone else who knew much about it in this country – I followed the following pattern – the first thing was that there was a fairly important paper published in a classical psychology journal – a high status mainstream journal which presented a theory of textual coherence and which was by a well known psychologist and linguist published quite recently then. I tried to find references to this in recent issues of *Science Citation Index* and *Social Science Citation Index*. I also went through the journals taken by the university library looking for ones which were relevant. I found there was one called *Discourse Analysis* which I immediately got hold of and read all the papers in. On the back of this journal it had advertisements from the publisher for related books – which I also got hold of – and that helped as well. One of the books was quite a good one – so I went through the abstracting journals again looking for references

to this person, and in each case of course, on finding a new journal I hadn't come across before if possible I would go through a few issues of that [MRC/SSRC 4A].

This pattern displays several features. First, identification or knowledge of a key reference. A very common way for social scientists to explore an area is to identify a key paper – either one which the social scientist already knows or is told of, one which is on the boundaries of his existing interests, or one which is found or worth following up from a keyword search. Second, the identification of references to this through *Science Citation Index* and *Social Science Citation Index*. Third, the identification of relevant journals – or more generally of relevant sources – from browsing in the library and through the journal papers. Fourth, working through the papers in those sources found to be relevant.

More formally these features may be described as representing the following characteristics:

Starting: identifying a key paper to commence the search;
Chaining: following up references to this paper, and following up book advertisements from the journal consulted;
Browsing: identifying relevant journal sources;
Extracting: working through material in relevant sources.

One of those interviewed in the Department of Psychology described how he kept up to date with material on the subject of memory. The pattern, or aspect of the pattern, described here, differs from the first, in that, in the first case, the starting point was a key paper or a key author which the researcher knew of, but on a topic or in an area with which he was unfamiliar, whereas in this case it was in relation to keeping up to date in an area where the individual was familiar both with the subject and the sources.

> Well the stuff on most things – the stuff on memory – there are various journals which I look through periodically. Every two or three months I go down to the library and have a look at the recent issues ... There are four or five journals that may well have something of note. The main source of information is *Current Contents*, which we get, the *Social Sciences Current Contents* and the *Life Sciences*. The *Life Sciences* has some stuff but it's usually not relevant – the *Social Sciences* is much more relevant to me ... Things like *Psychological Review* it depends very much who the editor is for the period of three years or whatever – the previous editor had quite a lot of stuff which I was interested in. *Psychological Review* covers a broad area and if it's interested in my area then I look at it a lot, whereas at the moment it's not so I don't have to look at it carefully. But there are things specialising in memory – *Memory and Cognition, Cognitive Psychology, Cognition* – a lot of those sort. All of which tend to have articles, some of them are slightly more hard-nosed than others, but there is a possibility that you'll get something in all of them. *Journal of Verbal Learning and Behaviour* ... and *Human Learning*

> is a new one that's come out quite recently, these are the ones that tend to be specialist in my area. But again there are the general ones like the *British Journal of Psychology*, and, of course, the *Journal of Experimental Psychology* has a section on human learning and memory ... I don't do a very thorough search of these I just wander down to the library and have a look through the recent articles [Psychology 4A].

This aspect of the individual's information seeking pattern displays two other significant characteristics of information seeking by academic social scientists:

Monitoring: maintaining awareness of developments in an area through regularly following particular sources;
Differentiating: employing differences in the nature of the source materials to filter material.

Several of the characteristics can be seen again in an aspect of the information seeking pattern of one of those interviewed from Education in relation both to material on a topic with which he was familiar, scripted drama for children, and one with which he was unfamiliar, television drama for children; for the latter topic he had had a computer based search carried out by the University library.

> You look at the publishers' lists, and you read through those it takes quite a long time ... I get the drama section of the Methuen output, for instance, Hutchinsons do a drama catalogue. They've got the Department on the list ... I get the drama quota. Sometimes I'm on the publishers' lists, with Methuen I think I am, Cambridge, who published my first book send me their catalogue, Macmillan send me theirs now automatically ... There are various journals that I've picked up articles from – *Use of English* was quite a fruitful field, *Theatre Quarterly*, which is now extinct, *Drama Quarterly*, which I believe is still going ... (The online search) was quite useful, hit on British and American books on the subject, I got quite a number of good books through that actually, mostly American ... It's a newer field for me than scripted drama where I could rely on a big backlog of experience, here I was feeling my way more, and wasn't sure, in fact, what had been written on the subject. I knew that very little was published in England and wanted to see what the Americans had done ... There are articles in British journals on the subject – things like *Two D, Dance and Drama*, I knew about *Two D* before, and things like the *Theatre Quarterly* which I knew about before, I searched through that and found articles that I hadn't realised were there before. It's only when you start looking for something that you find them [Education 1D].

This pattern displays characteristics relating to starting (the employment of an online search to locate references), differentiating (between journals and between publishers), monitoring (both publishers' lists and journals), and extracting (from the same publishers' lists and journals).

A final example, from one of those interviewed in Economic and Social History, illustrates the complexity of the information seeking patterns, and the way in which the different characteristics of the patterns interlink and complement each other. The information seeking activities described took place in the context of searching for information to set up a new course.

> Last year I had to set up one that was very much outside my normal range – teaching the industrial revolution period as history ... I started with the last five years' *Economic History Reviews*, which are standard, that is the economic historians' journal – *Economic History Review*. I looked through that for the last five years what articles there were on that period, what books had been published on that period over the last five years, used those as the basis for a literature search. Read around there in journals and books, and from bibliographies in those to bibliographies in other things ... And then there are the standard specialist things like the Newcomen Society for Engineering History, and the *Agricultural History Review* for agrarian history, and so on, and you go to these. It's the obvious ones to start with [Economic and Social History 4C].

This account includes starting (with a review type journal), extracting (employing the same journal), chaining, and differentiating (both general and specialist sources in the field).

Six characteristics seemed sufficient to subsume the different generic features of the various patterns, and, at the same time provide the framework for a flexible model to underpin recommendations for information retrieval system design and evaluation:

1. *Starting:* activities characteristic of the initial search for information;
2. *Chaining:* following chains of citations or other forms of referential connection between material;
3. *Browsing:* semi-directed searching in an area of potential interest;
4. *Differentiating:* using differences between sources as filters on the nature and quality of the material examined;
5. *Monitoring:* maintaining awareness of developments in a field through the monitoring of particular sources;
6. *Extracting:* systematically working through a particular source to locate material of interest.

The six features of the model together represent the major generic characteristics of the social scientists' individual information seeking patterns, and any individual pattern can, therefore, be described in terms of the features of the model. However, the detailed interrelation or interaction of the features in any individual information seeking pattern will depend on the unique circumstances of the information seeking activities of the person concerned at that particular point in time. The relationships between the features of the model can, therefore, only be indicated in the most abstract and general terms unless there is reference to a particular information seeking pattern.

For example, in general terms, relationships between features of the model can only be described hypothetically, i.e. starting may lead to chaining, differentiating may play a role in identifying sources for monitoring, or extracting may complement monitoring, but, in terms of any individual information seeking pattern (or particular sequence of information seeking activities), it would be possible to state these relationships categorically, i.e. starting did lead to chaining, differentiating did play a role in identifying sources for monitoring, and extracting did complement monitoring. The model does not, therefore, constitute a hierarchic sequence for classifying individual information seeking patterns, nor a prescriptive set of search heuristics, but rather a set of related categories which, taken together, can be used to describe individual information seeking patterns, and perhaps help to explain details of their topography.

3. THE BEHAVIOURAL MODEL AND INFORMATION RETRIEVAL SYSTEM DESIGN

In this section the features of the behavioural model are described in more detail, and the results of the analysis are compared both to research carried out on human-computer interaction in an information retrieval context, and to the extensive body of literature on social science user studies. Although some of the features of the model are present in existing operational or experimental systems, described below, the application of the model to questions concerning information retrieval system design here represents an attempt to provide a comprehensive analytical model from which recommendations for system design may be derived, rather than a set of detailed specifications for the design of an actual system.

3.1 Starting

Starting refers to characteristics of the information seeking patterns of researchers who are commencing work on a new topic or in a new area. The researcher may be experienced or inexperienced in research, and may have some or no familiarity with the topic or area. A variety of means were employed by those interviewed to obtain information when they were starting on new topics, including the use of informal contacts, the importance of which is frequently mentioned in reviews of social science information use [29–31, 38, 46, 49, 58, 59], and was equally prominent in the information seeking patterns of the social scientists interviewed for this study.

3.1.1 Starter references

The use of starter references whether previously collected or newly recommended allows the researcher not only to get some purchase on a new subject but also to do it in a way which allows other information gathering activities to be quickly established. They serve as a starting point into an area from which other ways of gathering material can be established. They are usually employed in two roles: to alert the individual to principal ideas or

key studies; and to provide overviews of an area and serve as the basis for chaining.

Many of those interviewed said that their first step in finding information on a new or relatively unfamiliar topic would be to seek out people who knew something about the area and ask them for references to introductory works, key references, and key authors. An advantage of this approach is that the contact typically provides evaluations of the quality or importance of the references provided. This allows the searcher to concentrate on what are perceived to be key references or key ideas [Psychology 6A: Continuing Education 10D: Economic and Social History 4C: Sociology 10C; 13C]. Students and novice researchers can be particularly dependent on supervisors or colleagues providing them with starter references, and those supervising research students or research assistants made explicit reference to this aspect of their supervisory role [MRC/SSRC 7A; 12A: Psychology 1A; 3A: Education 6D; 7C: Economics 1C: Economic and Social History 5C: Sociology 10C].

Most of the researchers were aware of a broader range of topics than those on which they were currently working, they often already knew of key people or key references in the new area, and their natural tendency was to begin with what they were aware of and commence the search from there. Some made explicit provision for this and maintained files of such references or made a note of references on topics tangential to their main interests – these then become potential starter references [MRC/SSRC 1A; 8A: Psychology 4A: Education 4D].

3.1.2 Reviews and review articles

Most subjects have some form of provision for the publication of reviews, whether this takes the form of annual reviews, or reviews in journals which also publish other types of articles. Some sources are recognised as being more likely than others to carry review type material, sometimes this is clearly indicated in the title of the source, in other cases those in the field know that a large number of reviews are published in a particular source [MRC/SSRC 12A: Economic and Social History 4C].

Reviews can be particularly valuable when starting since they not only represent a source of reference to primary material but also provide a context or framework for understanding that material. Reviews – and review type material such as synoptic tutorial articles – were particularly singled out by the psychologists interviewed as representing a convenient means of getting started in a new area [MRC/SSRC SAPU 1A; 4A; 8A; 5D; 7A: Psychology 4A].

Collections of papers were also perceived as providing useful overviews of an area, these having the advantage that it is the primary material itself which is being consulted avoiding, to an extent, the problem of secondary material distorting the primary literature [MRC/SSRC 11A; 3B].

3.1.3 Library catalogues, abstracts and indexes

An obvious way of starting in a new area is to consult bibliographies, abstracts, indexes, and library subject catalogues. Most of those interviewed

had undertaken a formal literature search of this kind, often when first embarking on their research careers, and some of the researchers considered that if they were to start in a new area they would undertake a search of this kind [Psychology 7A: Continuing Education 9D: Economic and Social History 5C: Geography 7C: Sociology 13C; 11D]. But the use of such services was not heavy and confirms the frequent observation of studies of the information seeking activities of social scientists that relatively low importance is attached to this means of locating information [38, 50, 60].

Online searching of databases represents a quick alternative to manual searching of secondary services, and all of those interviewed were asked if they had ever had an online search undertaken for them. Their opinion of the utility of such searches was mixed. In general there seemed to be greater satisfaction with the results of the search when the individual was unfamiliar with the area and was trying to find a quick way into the literature, particularly if that literature was diffuse [MRC/SSRC 5D: Continuing Education 10D: Education 2D; 1D].

Computer based searches were not, however, perceived as trouble free ways of information searching. Those who had had online searches carried out frequently underlined the difficulties which they had experienced – often making some attempt to explain those difficulties in terms of the limitations of the system. Whether the difficulties, in fact, derive from the nature of the systems used, from poor searches, terminological problems, or lack of suitable databases, it was clear that the use of online services was not seen as a sophisticated alternative to their other information seeking activities [MRC/SSRC 2D: Education 4D; 5D; 8D: Sociology 11D]. However, even in cases where the results of the search were not thought to have been satisfactory the search itself may still have proved useful. Either because the alternative would have been an exceedingly tedious and time consuming manual search, or because the search proved useful for identifying material from which the individual could proceed to follow other information seeking activities such as following chains of references [Education 6D: Sociology 11D].

3.1.4 Starting and information retrieval system design

It was clear from the comments of those who had had computer based searches undertaken that such searches were often undertaken at the start of a project, and with the intention of providing some references from which the individual could proceed to other information seeking activities, particularly chaining. Often the researcher starting out on a new topic is looking for something which can serve as the basis for chaining, and starting can change almost immediately to chaining as the individual moves out from the original key reference or references to material cited in that material [MRC/SSRC 12A; 5D; 7A; 3B: Psychology 1A: Economics 1C: Prehistory and Archaeology 15D: Sociology 11D; 14C].

This point is of particular significance for those concerned with information retrieval system design as searchers may come to a retrieval system with a collection of such references, in which case they might proceed directly to

chaining, or they may hope that the system will provide them with some starter references which can serve as the basis for chaining. In the former case it is probably desirable that the system be set up to 'prompt' the searcher for starter references. The database could then be searched to ascertain whether any of these are present or have been cited in other material on the database.

If the searcher has no such references, or wishes to identify more, the first step, obviously, would be to try to identify material which seems to match the terms of the subject description supplied by the searcher. If the searcher finds it difficult to discriminate between references obtained from a preliminary subject search then the system may be able to help the searcher select those which seem most likely to be useful in the two roles previously mentioned (i.e. to alert the individual to principal ideas or key studies; and to provide overviews of an area and serve as the basis for chaining). In the first category the searcher may be interested in seeing material which has been highly cited, either by the set of references obtained, or from that set. The fact that material is highly cited does not entail that it will definitely be of interest, but the searcher may wish to know whether any material is heavily cited in the area, if only to be aware of its existence.

In the second category, any material which itself has a large number of references is likely to be worth noting as a potentially rich source for chaining. Reviews, review papers, and synoptic articles can be particularly valuable in this role, as the searcher may hope to identify the more important theoretical developments and empirical studies from seeing them cited in reviews, and then following up citations to the individual papers. A facility to identify review type material, or material which made large numbers of references could aid the searcher to select references which were most likely to provide an overview of an area and be useful as sources of references for backward chaining.

Someone entering a new area may also find it useful to have some indication of the sources that publish material of interest. These could be identified either from the sources in which interesting items themselves appeared, perhaps arranged in order of the frequency in which interesting material appeared, or from sources which were cited in material of interest, again possibly ordered in terms of the frequency with which they were cited. The identification of potentially useful sources in this way could also provide the basis for browsing, differentiating, monitoring or extracting.

The conclusion of the first attempt to find references in a new area may represent only the beginning of the search activity. At a later stage other criteria involving more sophisticated judgements on the state, pre-occupations and methods of the field may modify the perceptions of the searcher. But this comes from building up a picture of the subject and from familiarity with the key ideas and people of the area, not from mechanically collecting references.

3.2 Chaining

Citation patterns and practices, and the significance of citation have been the subject of a very large number of studies covering a broad range of disciplines.

A comprehensive review of studies of citation practices, and of arguments concerning the role and significance of citation has been provided by Cronin [61, 62]. The focus of interest here, however, is not with citation practices, as such, but with characteristics of patterns of searching for information which involve following citation connections between material. This activity will be referred to here as chaining. Chaining can take two forms:

backward chaining – following up references or sources cited in material consulted;
forward chaining – identifying citations to material consulted or known.

Backward chaining is a traditional feature of information searching in all academic disciplines: its importance for the information seeking activities of social scientists has been underlined in a number of studies which have identified it as by far the most frequently mentioned aspect of formal information seeking, and often as the aspect of formal information seeking which is considered most important by social scientists [38, 63–65]. Forward chaining, except in the context of legal citation, is a relatively new innovation, and dependent on the use of special bibliographic tools in the form of citation indexes [66, 67].

3.2.1 Backward chaining

Following up references or footnotes was a major characteristic of the social scientists' information seeking patterns. All of those interviewed made some mention of it, and many employed it as their principal means of gathering information [MRC/ESRC 3B: Continuing Education 10D; 11C: Education 4D: Economics 1C; 3D: Prehistory and Archaeology 15D: Sociology 10C; 12C; 14C]. References in items identified from the original material can themselves be followed up, and their references, in turn, can serve as the starting point for further backward chaining. This can aggregate material very rapidly with a large number of references being accumulated through a sort of snowball effect [MRC/SSRC 12A: Psychology 2A; 9A: Politics 9C].

Chaining can be employed to pick up material in sources not followed, from references in material in sources followed, and so may be seen as complementing the monitoring activity, as monitoring a comparatively small number of sources – perhaps half a dozen to a dozen – can be relied on to lead elsewhere through chaining [MRC/SSRC 8A: Economic and Social History 5C: Politics 9C].

3.2.2 Forward chaining

Forward chaining, probably because it relies on the use of specialised bibliographical tools – the Institute for Scientific Information's various citation indexes – and is not a traditional part of information seeking practices in the social sciences, is less widely used and understood [Psychology 1A]. Nevertheless, forward chaining can represent a very effective means of checking whether further work has been done which has cited material already known of or consulted, but which has not come to the individual's notice

through his other information seeking activities [Psychology 1A]. The existence of the Institute for Scientific Information's various citation indexes, which now cover the sciences, social sciences and arts and humanities, enables forward chaining to be undertaken relatively simply, and sometimes with considerable success [MRC/SSRC 4A].

Although most of the social scientists were either unaware of the existence of these indexes, or of what could be done with them, a significant number of those who were aware of their existence had incorporated forward chaining into their information seeking activities. In this respect, the existence of these indexes may be said to represent an addition to the type of information seeking activities which may be undertaken by researchers [MRC/SSRC 9B; 10A: Psychology 4A: Economics 1C: Prehistory and Archaeology 15D: Sociology 13C].

3.2.3 Closure

The making of citations involves a considerable degree of subjectivity, and citations may be made for a variety of reasons [61, 62, 68–70]. Nor is it always the case that citations are considered important, or even particularly relevant by either the citing authors or the authors cited: in a study in the area of business administration less than a third of the citations made were considered essential by those citing [71], a phenomenon also noted in connection with author assessments of the relevance of their own cited references in the Cranfield II tests [72, 73], and remarked on by a number of those interviewed in this study [MRC/SSRC 4A: Psychology 6A; 2A].

The subjectivity applies equally to decisions to chase citations. Obviously, where the individual is immersed in the literature of the area many citations will be familiar [MRC/SSRC 11A: Psychology 6A: Economics 1C]. Decisions to follow up particular trails are related to the interests of the individual reading the article, and often the context of the citation is a key factor in deciding whether or not to chase it up [MRC/SSRC 7A: Psychology 4A; 3A: Education 5D: Prehistory and Archaeology 15D].

To a certain extent chaining may be intensified at the initial exploratory stage when an individual is entering a new area or starting off on a new topic and then again at the stage of completion when the individual wishes to ensure that he has covered all the important references and that nothing important has appeared which might have been missed by the other means by which he gathers information [MRC/SSRC 1A: Psychology 9A: Sociology 10C].

Sometimes a point is reached where the citations seem to be becoming more and more peripheral to the main subject of interest, or start dealing with increasingly minute aspects of it. Similarly, chasing up citations may draw to a close when the same references start appearing over and over again, or new references are adding little which is new to the picture the individual has built up, or are becoming more and more marginal. The citations may also be leading off into areas of literature that the individual does not really want to follow up because to do so would take him too far from his original concern. Other signs may be those of persistent corroboration or increasing specificity,

perhaps even a change of level or pre-occupation in the material cited [MRC/SSRC 12A; 2D: Psychology 9A: Sociology 11D].

The feeling of confidence that a search had been sufficiently thorough was described by one of the researchers as providing a sense of closure. This accords with the common meaning of the term which denotes closing, a closed state, or closing of a debate, and so the term 'closure' has been applied here to refer to the criteria mentioned as factors in the decision to bring citation chasing to a halt [MRC/SSRC 12A].

3.2.4 Chaining and information retrieval system design

Forward and backward chaining can be undertaken using online versions of the Institute for Scientific Information's citation indexes, but this requires precise keyboard specification at input, no abstracts are available, and the only form of subject access is via the terms occurring in the titles. As Morehead and Rouse observe, 'no readily available database provides a user with online access to articles, abstracts, text and reference lists' [74, p.196]. In the absence of an operational database which possesses these features, Morehead and Rouse's tests of information seeking behaviour where both citations and abstracts are available have had to be carried out in an artificial environment with an experimental system – the Data Base Access and Search Environment (DBASE) [74–77]. Using this experimental system enabled close examination of the human-computer interaction involved in searching it, but its limits place constraints on the type of searcher behaviour it is possible to simulate and study. In fact, the only types of searcher behaviour it is possible to study in the DBASE environment are 'brute force' searching where searchers browse through a large number titles of items on the database identifying a small percentage of these as relevant, and 'structured' searches where the searcher utilises citation relationships to identify relevant items [74].

A retrieval system designed for social scientists should provide facilities for both backward and forward chaining. At its simplest the system should enable the searcher to examine lists of references in material thought to be of interest, the decision as to which citations to follow, and the extent to which citation chains were followed, being left to the individual. Any references identified in this way, if they were available on the database, should also be searchable, in turn, for their references. The facility for identifying references which the searcher wishes to see, and those he wishes to follow up further to examine their lists of references, should be made as simple to operate as possible, perhaps by the employment of a movable cursor on a screen display.

In a similar way a facility for simple forward chaining should be available. In this case the searcher should be able to identify items which he would like to see references to, either from material which has been identified by some other form of searching on the database, or from material which the searcher knows of before commencing the search. Again, the facility for indicating which items the searcher wishes to see references to should be as simple as possible, the obvious method being to employ the same system feature as is used for backward chaining with a clear indication that it is forward chaining which is

taking place, perhaps by means of a screen display message to indicate whether the system was forward chaining or backward chaining.

Simple enhancements to the basic backward and forward chaining facility might be provided to aid the searcher. In terms of backward chaining the system could indicate that there were other papers on the database by the same author. This could be either the author of the paper whose references were being followed, or the authors of the references themselves. The searcher could then decide whether to examine these in addition to the items identified by chaining. Forward chaining might be broadened in the same way to include references to other papers by the same author, or to all references to any papers by that author. In either case the system should indicate to the searcher the level of specificity or generality of the chaining being undertaken, and, allow the searcher to limit or extend the level of generality as the search proceeds. This would enable the searcher to restrict a simple citation search if he felt too many, or marginal, references were being identified, and conversely, to broaden the search if too few, or no, references to a particular paper were found.

The employment of computer-based systems enables other, more sophisticated forms of citation chasing, such as bibliographic coupling or co-citation searching to be carried out without much extra effort on the part of the searcher. It would be desirable for the system to offer such facilities. However, as such forms of citation searching are not a conventional part of the information seeking activities of social scientists a brief explanation of the concept of bibliographic coupling or co-citation might be provided for those who require it. The existence of a bibliographic coupling or co-citation facility should also be made known to any searcher employing the chaining facility, perhaps via a screen menu. This would leave the individual researcher with the choice whether to follow individual cited papers or identify papers which were related in terms of bibliographic coupling or co-citation.

Following chains of citations is related to the interests and perceptions of searchers, who are usually only interested in following up a limited number of references, and social scientists are often very selective in the references they follow up, employing interpretive skills which it is difficult to conceive of modelling in an automatic process. Therefore, in most cases, the decision as to which citations to follow up should be left to the searcher – not pre-empted by an algorithm which mechanically followed up all the branches.

However, in some cases the searcher may wish to attempt to identify all the connecting chains which do exist; in such cases the provision of a facility for identifying all reference connections, up and down, to material of interest would be valuable. This would enable the searcher to be sure that, at least on the database searched and at the level of generality specified, he had exhausted all the possible trails, or identified all the existing citation connections. Such a facility might be particularly useful if other forms of chaining had produced a small number of references, if the searcher wanted to follow very long chains, or if the searcher wanted to be certain he had undertaken a comprehensive search.

To a certain extent any retrieval system (other than one providing full text) which offers a facility to follow chains of references has to place more emphasis on the act of deciding to follow up particular references rather than on the reasons underlying decisions to follow up particular references. In this respect, it must be admitted that any system which divorces citations from information about the citations, in terms of their context in the citing work places an extra step in the chaining activity, and breaks down an integrated action (seeing a reference in the context of an account or argument and deciding to follow that reference up) into a two stage procedure, with far less information being available for the searcher to make a decision than is the case where a reference is seen in its context in the citing article.

3.3 Browsing

The concept of browsing has several different connotations, a variety of different types of activities have been associated with it, and a number of different typologies of browsing have been put forward [78–82]. Nevertheless, whatever the connotation employed, browsing is a recognised, if sometimes disparaged form of information seeking activity, and, if purely random browsing is excluded, the major recognised forms of browsing can all be understood as representing forms of semi-directed or semi-structured searching in an area of potential interest. It is this connotation which will be followed here.

3.3.1 Semi-directed or semi-structured searching

Browsing, in the sense of semi-directed or semi-structured searching, in an area of potential interest, was an activity that many of those interviewed had engaged in at some time or another. It was also employed as a means of maintaining current awareness, either by scanning sets of recently published journals or *Current Contents*, or by examining recent book acquisitions [Education 4D: Economic and Social History 5C: Economics 1C].

Typical ways in which the social scientists browsed were by looking through the contents pages of journals in a broad subject area, checking the periodicals held by the library, or simply browsing along the shelves, either of journals or books. The pre-requisite for the activity to be effective, was that there should be at least some collocation of like material. The actual form of the material seemed less significant, in terms of its potential for browsing, than the fact that related material was grouped together [MRC/SSRC 9B: Education 8D; 7C: Geography 7C: Sociology 12C]. This is related to the fact that if browsing is to be effective some restriction has to be placed on the area of potential interest to be searched [MRC/SSRC 9B; 4A: Education 3C].

Abstracting services may be treated in a similar way. The searcher browses through issues of the abstracts not particularly looking for a particular item or for material on a particular, well defined, topic, but simply checking on the kind of work being carried out in an area or seeing if there is anything interesting which might be worth reading or following up [Education 3C: Economics 2D]. In the case of books, the obvious means of browsing are along

library shelves, the shelves of bookshops, and book displays at conferences [Education 8D: Sociology 14C].

In addition to its role in the identification of material, browsing can also serve the purpose of, or be directed towards, familiarising the researcher with the sources and material of an area. In this case, the activity can be seen to have two aspects, familiarisation and differentiation. Familiarisation allows the researcher to become aware of the sources of material in an area, of what there is available. Differentiation occurs as the researcher develops a knowledge of differences between the sources of material, that is, an appreciation of the differences between what there is available. So, although browsing and differentiating represent different characteristics of information seeking patterns, they can be related to each other in terms of the extent to which the social scientist differentiates between sources [Psychology 9A: Education 5D: Economics 1C].

3.4 Browsing and information retrieval system design

Conventional retrieval systems which have a 'browse' command are usually employing the term either to describe a facility to examine a set of references retrieved by a previous search, or to explore a portion of the indexing vocabulary online [83]. The ability to display retrieved sets of records online is a quite general feature of interactive online information retrieval services; decisions can then be made online to print only certain documents from the retrieved set. The ability to explore the indexing vocabulary is available on, amongst other hosts, Dialog using the 'expand' command, BRS employing the 'root' command, and Orbit using the 'neighbour' command [84].

In addition to these features of operational information services, a number of experimental systems have been proposed or set up which provide different degrees of provision for browsing [79]. An experimental approach to developing a browsing mechanism was undertaken in the context of the BLEND project. In this case browsing was simulated by providing the user with a detailed summary of the contents of a document and allowing him to 'jump' to any point in the document text and back out again [85].

Visual approaches to browsing have been put forward by Frei and Jauslin [86] using a graphical approach, and by Palay and Fox [87] in which the user would browse through a database by means of a series of screen menus. A large number of studies relating to graphic displays of thesauri have been undertaken [88, 89] including the experimental Pride retrieval system developed by Bovey and Brown [90] which employs graphics facilities to enhance user-system interaction, and integrates menu guidance with 'mouse' selection facilities. A related operational application of the visual approach is the employment of the ALS browser terminal in online library catalogues [91] which has been used as the basis of the Public Information in Rural Area Technology Experiment (PIRATE) [92] in which a database management system containing tourist and public information is operated by simple screen displays on the 'touch sensitive' screen.

The simplest way to incorporate a browsing facility into the system is to allow the searcher to browse, while the system provides the searcher with the information he requires to browse effectively. A retrieval system with a browsing facility should enable browsing to be carried out easily, and in all the forms of access which a searcher might consider useful. In practice, this means that any type of information held on the database should be directly accessible for browsing by the searcher. The principal types of information which the searcher is likely to wish to browse through are lists of authors, lists of journals, conference proceedings, or books, ideally with access to their contents pages, lists of cited works, subject terms, and broad subject headings.

Providing access for browsing for most of these types of information is relatively unproblematic, as the searcher is merely browsing through contents lists of material contained on the database. A facility for obtaining more information on an item in the list is more complex to set up in that it requires the existence of a link between items on the list and any information concerning those items. This requires that the database be constructed along lines that provide browsable lists of these various categories of information, and a simple means of moving between lists and obtaining further information on items in the lists. So that, for example, if a journal title seemed to be of interest the searcher could proceed to examine the contents pages of the journal, be provided with a list of articles in it, or citations to it.

IME's TinMan software represents a solution to this problem in that the searcher is able to 'Browse' and 'Navigate', following up connections between items on the database [93, 94]. Ayris describes how this facility to 'browse' and 'navigate' might be employed in the course of an actual search on a database:

> In its most advanced form, navigation allows the user to follow links between items in the database. If each set of items is so ordered that browsing access is possible on them in isolation and a navigational mechanism allows links to be followed, a search may be pursued ... For example, a user can start with a known author and find books by him, then link to the subjects of those books, then use the subjects to determine theses of interest, then use the authors of the theses as links to more recent articles [79, pp. 87–88].

A more general problem is raised by browsing as a form of subject access. Providing a list of subject terms in the database is similar to providing other lists of information concerning material in the database. However, searchers, when browsing manually, frequently employ the general classification or other form of shelf or subject organisation used in their library as a means of restricting or bounding the area in which they browse. If this facility were to be provided in an exploratory retrieval system some form of broader and narrower subject description would be necessary, with provision for the searcher to browse at a given level of generality or to change levels to make his search more general or more specific. The simplest way to effect this might be to provide a simple thesaurus structure, in addition to free text access to subject terms, and to employ the hierarchic structure of the thesaurus to

provide the means for the searcher to broaden or narrow his search. The thesaurus approach could also be used to provide related search terms for the searcher to employ in his search.

A browsing facility might be enhanced by the use of graphical displays of terms similar to that employed in I3R [24–26]. It may also be possible to enhance the facility by the provision of information derived from statistical analysis of the database, perhaps in terms of citation linkages. The information derived from such analyses would not, however, be used to determine where or how the searcher could browse, but could rather serve as an aid to the searcher in his browsing activities if he wished to employ them.

3.5 Differentiating

In most disciplines there are fairly well defined author and journal hierarchies, and different journals cater in a variety of ways for different specialisms, approaches and audiences. Garvey and Griffith [33] observed that psychologists had a very clear perception of the relative status of different journals, and of the different substantive and methodological orientations of particular journals. A number of studies have enumerated differences between journals according to the frequency with which they are cited [95–97], perceptions of their status, orientation, or quality [98], and on the rejection rates for publication [99].

Researchers frequently employ such differences between sources (or types of source) as a filter on the material examined. It is perhaps one of the rewards of specialisation that the researcher familiar with a field will be able to discriminate between sources in ways which may seem arcane to those less familiar with the field. Brittain described this aspect of the researcher's tacit knowledge of a discipline or field succinctly:

> To the outsider the problem of retrieving and collecting information about a particular topic may appear bewildering: faced with hundreds of journals, all of which could, with varying degrees of probability, contain the desired information, the problem of obtaining knowledge looks formidable indeed. To the researcher well versed in his field a number of sources appear as obvious first choices. For example, the strictly experimental psychologist has a number of 'core' journals which include those regarded as methodologically sound: he will turn to the *Journal of Experimental Psychology* or to the *Journal of Comparative and Physiological Psychology* (depending on interest) for details of experimental work, and to *Psychological Review* and *Psychological Bulletin* for state-of-the-art surveys on given topics [46, p.138].

Employing such differences between sources or types of source to filter material examined will be referred to here as differentiating.

Differentiating is effected by the researcher identifying different sets of sources in terms of the differing probability of their containing useful material. The manner in which differentiation between sources or types of source is effected, rather than differentiation in the fields themselves will be analysed

here. However, as differences between information sources were used extensively by those interviewed to delineate differences in their fields, differentiating between information sources can be employed as an indicator of differences in the subject.

For differentiating proper to be a significant element in a social scientist's information seeking pattern a reasonable awareness of potential sources, and a relatively sophisticated knowledge of the differences between these sources is required. This allows the social scientists to concentrate their information gathering activities on those sources which they perceive as having the highest likelihood of containing material which is relevant, at an appropriate level, and of the right type.

The criteria which seemed most significant for differentiating material were:

- the substantive topic of study;
- the approach or perspective adopted;
- and the quality, level, or type of treatment.

These differentiating features can be separated for purposes of analysis, but in practice they are often employed together. Furthermore, the same criterion can be used in different ways, for example, the level of treatment criterion could be employed to exclude material too specialised, detailed or technical for the requirements of one person, or too general, popular, journalistic, or lacking in rigour for the requirements of another.

3.5.1 Differentiating by substantive topic

The most obvious form of differentiating is in terms of the substantive topic of a source. Social scientists identify those sources which either focus specifically on a particular subject, or which regularly carry material on a particular topic [MRC/SSRC 1A; 2D; 8A: Psychology 8A; 2A; 4A: Continuing Education 10D: Education 5D; 6D: Economics 2D: Geography 6C: Politics 8C].

3.5.2 Differentiating by approach or perspective

Differences between various 'schools' of thought in the social sciences have been studied in the context of a number of subjects [46, 100–105]. Many of those interviewed made some reference to the importance of the particular approach or perspective of work in their assessment of relevance or interest. Often, if the social scientist is coming to a topic from a particular approach or with a particular perspective, pertinent material will appear in sources which have that approach or take that perspective. Distinguishing sources which were most likely to publish material having that approach or perspective served as a means of filtering the amount of material examined. This filtering could be based on differences in the disciplinary approach adopted, differences of conceptual interpretation, or differences in methodology [MRC/SSRC 2D; 9B: Psychology 7A: Education 2D; 3C].

3.5.3 Differentiating by quality, level, or type of treatment

Different periodicals in a field have different reputations, and there is

frequently a good perception of the relative prestige or quality of journals in an area. Differentiating between sources, therefore, may be employed, to assess the probable quality of the material [MRC/SSRC 10A; 4A; 5D: Psychology 2A]. Material which is apparently on the right topic may be treated at a level which makes that material inappropriate for the individual carrying out the search. This may be because the material is too general or too technical, not of sufficiently high quality, or addressed to the wrong type of audience. A particularly important distinction to note in this respect is that between academic and practitioner journals, some social scientists being interested in seeing both types of material, others concentrating on one or the other. In either case differentiating between sources can provide the individual with a useful filter on the level and type of material examined [MRC/SSRC 4A: Continuing Education 9D: Education 7C; 8D: Economic and Social History 5C; 4C: Politics 9C: Sociology 13C; 10C].

3.5.4 Differentiating and information retrieval system design

Differentiating is possible on conventional computer-based information retrieval services, for example, in some files on the Dialog system document types can be specified and searches limited to particular sources or types of documents, but this is implemented via the creation of a separate Boolean set of all the documents in the particular source or of the particular type, and requires precise keyboard specification at input. This makes the actual process of differentiating exceedingly complex to carry out and requires an exceptionally good knowledge both of the system and of the database being searched on the part of the searcher.

Conventional searches of computer based services, however, usually produce undifferentiated sets of references. No attempt is made to discriminate between material in terms of approach, perspective, level quality or type of treatment adopted. It is probably largely for this reason that such searches are often criticised for being insufficiently discriminating. The means by which social scientists differentiate between material on a topic in terms of differences of approach, perspective, methodology, quality, level or type of treatment create problems for conventional computer based information retrieval systems because such systems rely primarily on the terms of the subject description to create the retrieval sets. This tends to result in distinctions between material deriving from the other criteria not to be made, unless there are associated differences in the terms employed.

A system which has no facility for the searcher to distinguish between material in terms of differences of approach or level of quality of treatment will not be satisfactory for searchers who do wish to make such distinctions. But it is not necessary for the system to deal explicitly with such differences or to devise sophisticated ways of representing them in the database. Differences between the sources of material can be employed in this role. This makes use of existing differences in the literature of a field to differentiate the material in appropriate ways. The searcher should then be given the opportunity to specify those sources he thought most likely to contain material of interest to

him, and this information could be employed when searching. Such an approach also has the advantage that no particular intellectual effort is required when material is put on to the database.

Providing a facility for discriminating between, and indicating preferences for, sources is far easier than attempting to deal with differences of approach and level of treatment directly. The system would only need a facility for the searcher to differentiate easily between various sources or types of sources and to indicate preferences between them. This could take place before or during the search and might be effected in different ways. The searcher might provide a list of sources and preferences before the search started, or browse broad subject categories of sources held on the database indicating which sources or types of sources he would have most or least expectation of finding useful material in; alternatively, the searcher might discriminate between sources as the search progressed. If differentiation was effected by browsing through sources or by identifying sources at the time of the search then it would be desirable to have a facility for identifying the sources and preferences by means of a screen cursor, similar to that used when chaining.

There are three ways in which a facility for differentiating between sources could be used in a search:

- to restrict the search to a limited set of sources or types of source;
- to exclude certain sources or types of source from the search;
- to rank material by source or type of source in which it occurred.

The availability of the different ways of differentiating could be indicated by a screen menu display, and, again as with chaining, it would be important to indicate to the searcher which way of differentiating was being employed as the search progressed, again perhaps by means of a screen display. The facility for ranking material by source or type of source would allow the searcher to arrange the references retrieved in terms of his own perception of the likelihood of material being of interest.

The searcher who is familiar with an area may find it relatively easy to specify the sources or types of sources likely to contain material of the right approach, perspective, level and quality for him. In this case the obvious way to proceed is for the searcher to indicate which sources or types of sources he was most, or least, interested in and to use one of the differentiating categories listed above. The search could always be broadened out if it were thought to have proved too restrictive, either to widen the subject of the search while retaining the restriction on the sources, or to retain the same subject terms but extend the search in terms of the sources being searched.

If the searcher is unfamiliar with the area, or with the concept of differentiation, guidance will need to be provided on how it is possible to differentiate between sources of material and how this can be employed to vary the search result. If a subject search has already been undertaken, differentiation may be effected either by the searcher indicating items of highest interest, and the system then identifying other retrieved items from the same source, or by the searcher separately specifying sources, types of source,

or other differentiating criteria, which could then be applied to the retrieved set.

It is not being assumed that in every case a researcher will want to restrict the search to a limited type of source material. In the case of some of those interviewed differentiating a limited number of sources was not possible or was very difficult. The nature of their focus of interests was such that a large number of sources carried relevant or interesting material, and the individual could not be confident that this material would be picked up from a limited set of sources [MRC/SSRC 3B: Psychology 9A: Economics 1C: Economic and Social History 5C]. In the case of two of those interviewed, the differentiation of material in terms of different approaches or perspectives was noted, but was not particularly useful, or made information searching harder, particularly in the case where information was required which drew from the different approaches or perspectives [MRC/SSRC 10A: Prehistory and Archaeology 15D].

Therefore, the searcher should be able to carry out a comprehensive search of the full range of sources – particularly if there is little material and what there is is widely scattered. However, if it is thought desirable to restrict the search to specified sources or types of source, then this option should be available; the intention should be to give the searcher the choice of options, not preclude them.

3.6 Monitoring

The continuous monitoring of developments in a field of study was an important part of the information seeking activities of many of those interviewed. The principal ways in which those interviewed for this study monitored developments in their fields were through the use of informal contacts, monitoring services and research directories, journals or newspapers, and publishers' catalogues. These different ways of keeping up to date were not mutually exclusive. Individuals would frequently monitor journals and rely on colleagues or associates to bring other material to their attention, or attempt to monitor both journals and publishers' catalogues as they appeared. Nevertheless, for the purposes of analysis these different ways of monitoring can be treated separately.

Someone who is not moving into an entirely new or unfamiliar field may have a very good idea of the types of material which are likely to be useful [Politics 8C] and an individual who has been working in an area a considerable time will be familiar with the existing sources in the area [Geography 6C]. In some cases, where the social scientist has been working in the same area for a considerable period of time, the present work may be based on a history of previous work done by the researcher himself [Psychology 6A].

Sometimes, a social scientist will monitor a small number of sources very carefully and expect to find interesting material appear in them fairly often – compared to a larger number that might be monitored to some extent but far less carefully or less frequently and in which the probability of something coming up is correspondingly less [MRC/SSRC 1A]. Alternatively, someone

might feel the need to follow up a large amount of material relating to his research interests but only be interested in following up particularly interesting or significant material relating to his more general interests. The monitoring of sources central to the individual's focus of interests would then be more concentrated and directed than the monitoring of areas which the individual may feel he has to keep up with in more general terms – say for teaching purposes [Sociology 10C]. The individual can then differentiate the sources for material in terms of the likelihood of them coming up with useful or interesting material, and, if interests change, the sources monitored may change or some of those monitored be monitored more closely [MRC/SSRC 8A].

The alerting function of citations reduces the need to monitor all the sources that might conceivably carry material of interest, as do things such as reviews, and publicity for books and conferences [MRC/SSRC 8A: Continuing Education 11C: Education 4D]. Specialist research directories or bulletins of research in progress can also serve as a convenient means of keeping abreast of what is taking place in the forefront of an area [Geography 6C: Sociology 14C].

3.6.1 Informal contact

The use of informal means of communication has been studied in a wide variety of contexts ranging from science and technology through to the social sciences and humanities. A number of different concepts have been used to describe or explore the role of informal contacts in communication, the most notable being that of the 'invisible college' [106, 107], and that of the 'gatekeeper' [108, 109, 110]. Both these concepts have been the subject of considerable interest and a substantial body of research on the contribution of 'invisible colleges' and 'gatekeepers' to information exchange, with particular reference to studies relating to the social sciences, has been reviewed by Cronin [59].

Many of those interviewed used informal contacts to help them keep up to date, some relied very heavily on such informal contacts to keep them abreast of developments, and others stressed the importance of such contacts. One of those interviewed employed the services of other individuals and organisations monitoring the field, using these to pre-select sources and material of interest [Politics 8C]. The Director of the MRC/SSRC Social and Applied Psychology Unit appeared to be operating as a semi-formal gatekeeper for the Unit [MRC/SSRC 3B; 5D].

Social scientists immersed in an area and familiar with others working in the area often rely on such contacts to bring news and information to their notice, and in this way they keep each other up to date. The more an individual becomes immersed in an area the more important the informal network may become to his total information seeking or information gathering activities [Psychology 1A; 6A: Politics 9C: Sociology 10C].

Information obtained through informal networks may then be integrated into the individual's overall information seeking pattern – employing informal

and semi-formal channels to complement or to form the major part of his information seeking or information gathering activities. This may be through contact with others concerned with the same subject or working in the same area, or through contacts with immediate colleagues. Where other individuals or organisations are applying very similar criteria to the source material as the individual himself would, he may rely on these individuals or organisations to filter material for him [Education 3C: Politics 8C: Sociology 10C].

3.6.2 Monitoring journals

Skelton [63, 64] observed that scientists were more likely than social scientists to use journal sources, while social scientists tended to use monograph and journal sources to an equal extent. From the results of the interviews with the social scientists in this study journals and monographs were used in approximately equal proportions for keeping up to date, but there were marked differences between individuals and subjects, with journals being predominant in some areas and monographs in others, and some individuals using both to keep up to date while others concentrated on one or the other.

A principal way in which journal monitoring is carried out is through the identification of a set of journals which seem to publish material of interest frequently. In order for this to be effective some restriction has to be made on the number of sources monitored through the prior differentiation of a range of journals within a field. Those who have been working in an area for some time have usually identified a number of sources which regularly or consistently carry material central to their concerns. This form of monitoring has two aspects – maintaining direct awareness of a limited number of sources; amd indirect awareness of the existence of other sources, and of material in them, from references in the sources directly monitored [MRC/SSRC 12A: Psychology 2A: Continuing Education 11C; 9D: Education 5D; 7C: Economics 1C: Geography 6C: Politics 8C: Sociology 13C]. In a similar way, some social scientists make considerable use of the press, in particular the 'quality' press, to alert them to such things as the existence of economic research reports, or policy decisions [Education 3C: Economic and Social History 4C: Politics 9C].

The use of *Current Contents* as a means of monitoring journals was particularly remarked on by the psychologists interviewed; both those in the MRC/SSRC Social and Applied Psychology Unit and those in the Psychology Department made considerable use of the service. In contrast, *Current Contents* was hardly used at all by the educationalists and other social scientists interviewed, although a couple did mention making use of other, similar, types of current awareness service [MRC/SSRC 7A; 10A; 4A: Psychology 6A; 4A; 5A; 7A; 8A; 9A; 1A: Education 8D: Economics 1C].

3.6.3 Monitoring material published in book form

Monitoring what is being published in book form can take place in a number of ways, the principal ones being regularly scanning publishers' lists, regularly consulting reviews or continuing bibliographies, and by checking new

accessions to the library. Examples of all these forms of keeping up to date with the monograph literature were mentioned by social scientists from the different groups, although it did not seem to be a particularly significant aspect of the monitoring activities of the psychologists interviewed either in the MRC/SSRC Unit or in the Psychology Department [MRC/SSRC 8A: Continuing Education 10D: Education 2D; 6D: Economic and Social History 4C; 5C: Geography 6C: Sociology 12C].

3.6.4 Monitoring and information retrieval system design

Facilities for monitoring in conventional computer-based services are generally restricted to current awareness searching. Only recently has there been provision in computer-based systems for searchers to monitor easily the complete contents pages of sets of journals using the Institute for Scientific Information's *Current Contents Search* on BRS [111]. There is no equivalent facility, however, for online monitoring of publishers' lists.

Implementing a monitoring facility in a retrieval system is straightforward. All that is required is that the searcher specify which sources he wishes to monitor and these sources can then be searched automatically. The search may be implemented each time the searcher uses the system, or, alternatively, each time the particular sources are updated on the database. The searcher could be informed that the sources he is monitoring have been updated when he enters the system, and asked if he wishes to carry out a stored search in them, or wishes, for example, to examine the contents pages of recent journals, or recent publishers' lists, or to check the titles of books added to the database in his area. The monitoring facility may then be used to carry out a specific current awareness search in all, or a limited set of sources, or to examine the contents of material of interest.

Searchers should be able to monitor different types of source material, books, journals, conference proceedings, or publishers' lists, and should be allowed to carry out a monitoring or updating search in all of these categories or to restrict the search to particular categories, or individual items within categories. The searcher could broaden or restrict the scope of his monitoring, perhaps depending on the amount of material identified at any particular time.

A feature of monitoring mentioned by a number of those interviewed, particularly those who made use of the *Current Contents* service was differential monitoring: that is that they would rank the sources they monitored in order of the likelihood of their containing pertinent material, and would concentrate their attention on those they ranked highest. This feature of the monitoring activity could be implemented in the retrieval system by ordering sources examined, or by ranking material identified in a current awareness search by the source in which the material appeared.

The secondary characteristics of the monitoring activity, keeping aware of other sources and of the material in them from references in material in the sources directly monitored, would be covered through the chaining facility. However, it might be desirable to enhance the monitoring facility by an

alerting function, so that frequently cited sources, or sources which had not been cited in previous updates of the monitored sources, but which were cited in the present search, were brought to the attention of the searcher, who could then decide whether to examine these sources directly, and perhaps whether to include them in his monitoring profile. This facility might be particularly useful where new journals appeared in the area, of which the searcher was not aware, but which were being cited in articles published in sources he was monitoring.

3.7 Extracting

Extracting refers to the activity of going through a particular source selectively identifying relevant material from that source. The source may consist of a run of a periodical, a set of conference proceedings, a series of monographs, the contents of an archive, a collection of publishers' catalogues, or bibliographies, indexes, or abstracts, whether continuing or closed. The activity usually requires the setting aside of discrete – and sometimes considerable – periods of time for working through the source. It is one of the most directed and focussed of information seeking activities.

There is a close relationship between monitoring and extracting. For example, if, for whatever reason, the monitoring of a source lapses this may be made up for by a retrospective extracting exercise in that same source. Similarly, monitoring a journal for current awareness purposes may be complemented by undertaking a retrospective extracting exercise for a particular topic within the same journal. The similarity of form between the two activities is modified by the fact that extracting is generally more concentrated and directed than monitoring, and there is a tendency for the social scientist who is extracting from a source to make more use of indexes to the source than when monitoring the same source. However, reservations were expressed about relying solely on indexes as indicators of the actual content of the sources, and, in addition to employing the cumulative indexes, the individual would attempt to make a direct examination of the sources themselves [Psychology 2A: Continuing Education 9D: Prehistory and Archaeology 15D].

The identification of potentially useful sources of material is a critical prerequisite to extracting from those sources. If a key source, or key sources can be identified then the source or sources may provide the basis for much of the social scientists' information seeking activities. Identifying possible sources may take place in a number of ways. The source may be recommended by a colleague or supervisor or may be accepted as standard for the field. Citations to a particular source may suggest to the individual that the source itself is worth examining for the purposes of extracting [MRC/SSRC 4A; 8A]. The incentive to proceed to examine a particular journal from seeing references to material in it is likely to increase with the number of references to the same journal [Education 8D]. Sources may also be identified through browsing [Education 8D].

Having identified a source the individual works through a sequence of it, directly or using whatever indexes are available, identifying relevant or

interesting material. This can be made easier if the sequence of the source is held in one place, if there are good indexes, particularly cumulative indexes if a large scale retrospective search is being undertaken. The criteria for what is of relevance or interest may be very clear beforehand, in which case the search will be very highly directed, or they may be more open, perhaps developing in the course of going through the source.

3.7.1 Extracting from journals

Extracting from journals can represent a very effective means of quickly identifying material on a topic, especially if the source used is perceived as standard for a particular field. Several of those interviewed commented on the importance of such standard journals, whether these were general to a large area or ones which were understood to be standard to a particular sub-aspect of a field. [Psychology 9A: Education 8D; 6D: Continuing Education 9D: Economic and Social History 4C: Economics 3D: Geography 6C; 7C: Politics 9C: Prehistory and Archaeology 15D].

Some of those interviewed had identified a number of sources which were of particular interest and had set aside a regular time for extracting from them [Education 8D: Economics 3D: Geography 7C: Politics 8C]. It is interesting to note that the criterion of relevance was not exclusively applied to individual articles but rather to the sources themselves, and that the identification of sources of interesting material is seen as at least as important as the identification of items of interesting material. Moreover, sources not considered relevant to an individual's interests may become relevant or important not only if the individual's interests change, but also if the editorial policy of the source alters, or if, for some other reason, more relevant material starts appearing in a particular source [MRC/SSRC 2D: Psychology 4A].

3.7.2 Extracting from publishers' catalogues

The use of publishers' catalogues for monitoring publications in book form has its complement in extracting. For many of those interviewed publishers' catalogues were a primary source of references and identifying material through them a major aspect of their information seeking activities. The means of obtaining the catalogues differed between individuals. Some employed the International Book Information Service (IBIS) as a means of preselecting those catalogues which they wished to see. Others had made direct contact with the publishers or bookshops [Education 7C; 1D; 2D: Economic and Social History 5C: Economics 2D: Politics 8C: Sociology 10C; 11D].

3.7.3 Extracting from bibliographies, indexes and abstracts

Although not representing a major aspect of the information seeking activities of those interviewed a number did extract from bibliographies, abstracts and indexes. The availability of a subject bibliography on the particular topic, or the existence of an abstracting or indexing service with exactly the right focus for the individual concerned represent obvious criteria for a social scientist to

employ these kinds of sources [MRC/SSRC 5D: Psychology 6A: Prehistory and Archaeology 15D: Politics 8C].

3.7.4 Extracting and information retrieval system design

Only recently has provision been made for searchers to proceed through sets of contents pages of a given journal online, and, again, this is through use of the Institute of Scientific Information's *Current Contents Search* on BRS [111], and, as with monitoring there is no equivalent facility for extracting online from publishers' lists.

Extracting can take place either by working through individual runs of journals, sets of publishers' lists, bibliographies, indexes or abstracts, or by consulting cumulative indexes to such sources. Both types of approach should be provided for in a retrieval system as some searchers prefer examining the contents of individual issues – either because they are not sure of their requirements or these are developing as the search proceeds, or because they do not have confidence in the indexing facilities – others prefer the convenience of employing a cumulative index.

For the first approach to be possible the system must be designed to allow continuous movement through different source streams, whether these be runs of journals, publishers' lists, sets of conference proceedings, laboratory memoranda, working papers, or whatever other form of material held on the database. This approach to extracting requires that the sources be recomposed more or less in their original form for searching purposes. For example, if a searcher wished to work through the contents of a particular journal over a period of years, employing the first approach, he would wish to examine the contents pages of individual copies of the journal. This is not a particularly complex feature to incorporate in a retrieval system; all that is required is that individual articles in any issue of the journal can be identified and presented together to the searcher, who could then proceed to the next issue of the journal, and so on until the end of the run, or period of interest had been reached.

The second approach to extracting involves the use of a cumulative index to identify material in a source or sources. This requires the provision an index to the material in the sources. If a hybrid approach to indexing is applied this index might be made up of terms taken free text from titles, or, if available, abstracts, supplemented by the provision of index terms from a thesaurus or other form of controlled vocabulary. This represents a conventional approach to computer-based subject retrieval, but it would be important to allow the searcher to limit the search easily to those sources he was particularly interested in extracting from, and to allow the searcher to broaden the search to bring in other sources if desired.

4. IMPLEMENTING THE BEHAVIOURAL MODEL ON AN EXPERIMENTAL HYPERTEXT SYSTEM

One approach to implementing the features of the behavioural model on an experimental system would be to employ standard hypertext software [112,

113]. There is a precedent for this kind of approach to database design for reference retrieval in the hypercatalog project proposals [114]. The database would consist of bibliographic descriptions of items with their references, the individual items being connected by citation and other types of link. Works by the same author or in the same journal, or citations to a particular work or author, could be identified by links, while other forms of presentation such as contents pages of journals, or publishers' lists, could be provided for in separate linked nodes.

The kind of relationships which it would be necessary to specify in a general reference system based on the behavioural model are similar to those outlined by Moon [115] for a personal information system:

1. that articles are contained in the same book, conference proceeding, or special issue of a journal;
2. citation linkages to or from articles;
3. subject connections not covered by 1 or 2 above;
4. that articles are by the same person, group, or organisation;
5. connections of articles with previously published work or republications;
6. user-defined relationships.

In addition, Moon [115] noted that it would be useful to provide facilities for the user to:

1. indicate the document type and its likely use for the user (that it is to be used in a particular article perhaps);
2. attach comments to the links.

Provision for these sorts of relationships and facilities would enable a searcher to interact more flexibly with the database than is possible with conventional systems and would allow a searcher more nearly to follow the kind of information seeking pattern which has the characteristics of the model outlined above. Their enhancement by other kinds of search aids such as information on citation structures, collection hierarchies, link structures, classification schedules or thesauri, as suggested in the hypercatalog proposals [114], would mean that the searcher should be able to follow familiar search patterns more effectively and with more confidence that the search had been thorough, and experiment with more adventurous ways of searching which use of the system should make possible.

5. CONCLUSION

The importance of user considerations in the design of information retrieval systems has been recognised for a long time. There is a substantial literature on human-computer interaction which is relevant to information retrieval research [116–124], human factors research played an important role in the development and implementation of the interfaces of commercial computer-based information services [117, 121, 125–127], and numerous other studies

have either provided guidance as to how user reaction might be employed in the design of computer-based information retrieval systems, or have been considered in the design of operational systems [124, 128–132].

However, although a considerable amount of research has addressed issues connected with interface design there has been less research directed towards understanding the nature of the interaction with the databases themselves. This is a significant deficiency, as the development of a comprehensive understanding of the nature of the human-computer interaction taking place with such systems requires understanding of the nature of the interaction with the database as well as of the interaction with the interface to the database. This is particularly important in relation to the interaction with hypertext databases, where there is more flexibility in the interaction of the user with the database, and the user's need for guidance in effectively navigating the database may be correspondingly greater.

It cannot be claimed that a model developed in relation to the information seeking activities of academic researchers in the social sciences can be employed to provide detailed guidance for the design of information retrieval systems for other types of users. Nevertheless, the general principle of using the behavioural aspects of users' information seeking activities to inform the design of information retrieval systems (which is a commonplace of the user studies literature) is more widely applicable, and could play a more prominent role in the design of computer based information retrieval systems than, at present, it does.

The model may also be of more general interest in that it highlights key features of the perceptions of academic social scientists concerning their generation, communication and utilisation of information. In this respect, the model may be seen to represent an empirically derived framework for further research into the academic communication process. A small study has been undertaken in relation to academics in the field of English literature [133] and it is hoped to carry out further studies both to assess the more general usefulness of the model for analysing the activities of researchers in different fields and also to refine the features of the model itself.

ACKNOWLEDGEMENTS

A number of people have been good enough to have given their time to talk to me personally about my work. My particular thanks must go to Nicholas Belkin, Cyril Cleverdon, Thomas Green, Christopher Needham, Stephen Robertson and Vi Winn.

I should also like to record my gratitude to all those social scientists who agreed to co-operate with me in the course of this study, and to David Jones of the University of Sheffield Library for providing me with records of online searches he had carried out in the library for social scientists.

For their criticisms, advice and encouragement I owe a particular debt of thanks to my colleagues in the Department of Information Studies especially Norman Roberts, Peter Willett and Tom Wilson.

Finally, I should like to acknowledge the support of the Department of Education and Science for providing a research studentship.

REFERENCES

1. SPARCK JONES, K. The Cranfield tests. *In:* SPARCK JONES, K., *ed. Information retrieval experiment.* London: Butterworth, 1981, 256–284.
2. ELLIS, D. The effectiveness of information retrieval systems: the need for improved explanatory frameworks. *Social Science Information Studies, 4,* 1984, 261–272.
3. ELLIS, D. Theory and explanation in information retrieval research. *Journal of Information Science, 8,* 1984, 25–38.
4. ROBERTSON, S.E. The methodology of information retrieval experiment. *In:* SPARCK JONES, K., *ed. Information retrieval experiment.* London: Butterworth, 1981, 9–31.
5. SPARCK JONES, K. Retrieval system tests 1958–1978. *In:* SPARCK JONES, K., *ed. Information retrieval experiment.* London: Butterworth, 1981, 213–255.
6. ODDY, R.N. Laboratory tests: automatic systems. *In:* SPARCK JONES, K., *ed. Information retrieval experiment.* London: Butterworth, 1981, 156–178.
7. DANIELS, P.J. Cognitive models in information retrieval – an evaluative review. *Journal of Documentation, 42,* 1986, 272–304.
8. ODDY, R.N. *Reference retrieval based on user induced dynamic clustering.* PhD Thesis, University of Newcastle upon Tyne, 1975.
9. ODDY, R.N. Information retrieval through man-machine dialogue. *Journal of Documentation, 33,* 1977, 1–14.
10. ODDY, R.N. Retrieving references by dialogue rather than by query formulation. *Journal of Informatics, 1,* 1977, 37–53.
11. BELKIN, N.J. *A concept of information for information science.* PhD Thesis, University of London, 1977.
12. BELKIN, N.J. Progress in documentation: information concepts for information science. *Journal of Documentation, 34,* 1978, 55–85.
13. BELKIN, N.J. Anomalous states of knowledge as the basis for information retrieval. *Canadian Journal of Information Science, 5,* 1980, 133–143.
14. BELKIN, N.J., BROOKS, H. *and* ODDY, R.N. Representing and classifying anomalous states of knowledge. *In:* MACCAFFERTY, M. *and* GRAY, K., *eds. The analysis of meaning: Informatics 5.* London: Aslib, 1979, 227–238.
15. BELKIN, N.J., ODDY, R.N. *and* BROOKS, H.M. ASK for information retrieval I: background and theory. *Journal of Documentation, 38,* 1982, 61–71.
16. BELKIN, N.J., ODDY, R.N. *and* BROOKS, H.M. ASK for information retrieval II: results of a design study. *Journal of Documentation, 38,* 1982, 145–164.
17. RICH, E. *Building and exploiting user models.* PhD Thesis, Computer Science Department, Carnegie-Mellon University, 1979.
18. RICH, E. User modelling via stereotypes. *Cognitive Science, 3,* 1979, 329–354.
19. RICH, E. Users are individuals: individualising user models. *International Journal of Man-Machine Studies, 18,* 1983, 199–214.
20. BELKIN, N.J., SEEGER, T. *and* WERSIG, G. Distributed expert problem treatment as a model for information system analysis and design. *Journal of Information Science, 5,* 1983, 153–167.
21. BELKIN, N.J., HENNINGS, R.D. *and* SEEGER, T. Simulation of a distributed expert-based information provision mechanism. *Information Technology: Research, Developments, Applications, 3,* 1984, 122–141.
22. BROOKS, H.M., DANIELS, P.J. *and* BELKIN, N.J. Problem description and user

models: developing an intelligent interface for document retrieval systems. *In: Advances in intelligent retrieval: Informatics 8*. London: Aslib, 1985, 191–214.
23. BROOKS, H.M., DANIELS, P.J. *and* BELKIN, N.J. Research on information provision mechanisms. *Journal of Information Science, 12*, 1986, 37–44.
24. CROFT, W.B. User-specified domain knowledge for document retrieval. *In:* RABBITTI, F., ed. *Proceedings of the 9th International Conference on Research and Development in Information Retrieval, Pisa, Italy, 1986*. Washington: ACM, 1986, 201–206.
25. CROFT, W.B. Approaches to intelligent retrieval. *Information Processing and Management, 23*, 1987, 249–254.
26. CROFT, W.B. *and* THOMPSON, R.H. I3R: a new approach to the design of document retrieval systems. *Journal of the American Society for Information Science, 38*, 1987, 389–404.
27. ELLIS, D. *The derivation of a behavioural model for information retrieval system design*. PhD Thesis, University of Sheffield, 1987.
28. ELLIS, D. A behavioural model for information retrieval system design. *Journal of Information Science*, 1989, to appear.
29. AMERICAN PSYCHOLOGICAL ASSOCIATION. *Report on scientific information exchange in psychology*. Volume 1. Washington, DC: American Psychological Association, 1963.
30. AMERICAN PSYCHOLOGICAL ASSOCIATION. *Report on scientific information exchange in psychology*. Volume 2. Washington, DC: American Psychological Association, 1965.
31. AMERICAN PSYCHOLOGICAL ASSOCIATION. *Report on scientific information exchange in psychology*. Volume 3. Washington, DC: American Psychological Association, 1969.
32. BATH UNIVERSITY. *Toward the improvement of social science information systems: overview of research carried out 1971–1975*. Bath: Bath University Library, 1980. (Design of Information Systems in the Social Sciences (DISISS) Research Report, Series A, no. 1)
33. GARVEY, W.D. *and* GRIFFITH, B.C. Communication and information processing within scientific disciplines: empirical findings for psychology. *Information Storage and Retrieval, 8*, 1972, 123–136.
34. GARVEY, W.D., LIN, N. *and* NELSON, C.E. Communication in the physical and social sciences. *Science, 170*, 1970, 1166–1173.
35. GARVEY, W.D., LIN, N. *and* NELSON, C.E. A comparison of scientific communication behaviour of social and physical scientists. *International Social Science Journal, 23*, 1971, 256–272.
36. GARVEY, W.D., LIN, N., NELSON, C.E. *and* TOMITA, K. Research studies in patterns of scientific communication: I. General description of research program. *Information Storage and Retrieval, 8*, 1972, 111–122.
37. GARVEY, W.D., GOTTFREDSON, S.D. *and* SIMMONS, J.G. A comparison of two major scientific information exchange processes in psychology: 1962 and 1976. *American Psychologist, 39*, 1984, 11–21.
38. LINE, M.B. The information uses and needs of social scientists: an overview of INFROSS. *Aslib Proceedings, 23*, 1971, 412–434.
39. SWIFT, D.F., WINN, V.A. *and* BRAMER, D.A. *A case study in indexing and classification in the sociology of education*. (Report for the period September 1970–June 1973, 2 vols.). Milton Keynes: Open University, 1974.
40. SWIFT, D.F., WINN, V.A. *and* BRAMER, D.A. A multi-modal approach to indexing and classification. *International Classification, 4*, 1977, 90–94.
41. SWIFT, D.F., WINN, V.A. *and* BRAMER, D.A. Multi modality in indexing and searching. *Journal of Informatics, 1*, 1977, 91–95.

42. SWIFT, D.F., WINN, V.A. *and* BRAMER, D.A. A sociological approach to the design of information systems. *Journal of the American Society for Information Science, 30*, 1979, 215–223.
43. WINN, V.A. A case study in the problem of information processing in the social science field: the OSTI-SEA project. *Aslib Proceedings, 23*, 1971, 76–88.
44. ADAM, R. Social science information and its users. *Journal of Librarianship, 3*, 1971, 150–157.
45. ADAM, R. Language and information retrieval in the social sciences. *Aslib Proceedings, 34*, 1982, 394–405.
46. BRITTAIN, J.M. *Information and its users: a review with special reference to the social sciences*. Bath: Bath University Press, 1970.
47. BRITTAIN, J.M. Information services and the structure of knowledge in the social sciences. *International Social Science Journal, 31*, 1979, 711–728.
48. ELLIS, D. Social science information research. *Journal of the American Society for Information Science, 37*, 1986, 86–88.
49. HOGEWEG-DE HAART, H.P. Characteristics of social science information: a selective review of the literature: part I. *Social Science Information Studies, 3*, 1983, 147–164.
50. HOGEWEG-DE HAART, H.P. Characteristics of social science information: a selective review of the literature: part II. *Social Science Information Studies, 4*, 1984, 15–30.
51. ROBERTS, S.A. Developing a focus for library and information research in social sciences in the United Kingdom. *Education Libraries Bulletin, 24*, 1981, 1–19.
52. GLASER, B.G. *and* STRAUSS, A.L. *The discovery of grounded theory: strategies for qualitative research*. Chicago: Aldine, 1967.
53. PATTON, M.Q. *Qualitative evaluation methods*. London: Sage, 1980.
54. GLASER, B.G. *Theoretical sensitivity: advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press, 1978.
55. TURNER, B.A. Some practical aspects of qualitative data analysis: one way of organising the cognitive processes associated with the generation of grounded theory. *Quantity and Quality, 15*, 1981, 225–247.
56. DIESING, P. *Patterns of discovery in the social sciences*. London: Routledge & Kegan Paul, 1972.
57. KAPLAN, A. *The conduct of inquiry*. San Francisco: Chandler, 1964.
58. ADAM, R. Can the transmission of sociological knowledge be made more effective? *International Social Science Journal, 34*, 1982, 329–345.
59. CRONIN, B. Invisible colleges and information transfer: a review and commentary with particular reference to the social sciences. *Journal of Documentation, 38*, 1982, 212–236.
60. BATH UNIVERSITY. *Experimental information service in the social sciences 1969–1971: final report*. Bath: Bath University Library, 1972.
61. CRONIN, B. The need for a theory of citing. *Journal of Documentation, 37*, 1981, 16–24.
62. CRONIN, B. *The citation process: the role and significance of citations in scientific communication*. London: Taylor Graham, 1984.
63. SKELTON, B. *Comparison of results of science user studies with investigation into information requirements of the social sciences*. Bath: Bath University Library, 1971. (DISISS working paper no 1)
64. SKELTON, B. Scientists and social scientists as information users: a comparison of results of science user studies with the investigation into information requirements of the social sciences. *Journal of Librarianship, 5*, 1973, 138–156.
65. STENSTROM, P. *and* MCBRIDE, R.B. Serial use by social science faculty: a survey. *College and Research Libraries, 40*, 1979, 426–431.

66. GARFIELD, E. Citation indexes in sociological and historical research. *American Documentation, 14,* 1963, 289–291.
67. GARFIELD, E. Citation indexing: a natural science literature retrieval system for the social sciences. *American Behavioral Scientist, 7,* 1964, 58–61.
68. MARTYN, J. Citation analysis. *Journal of Documentation, 31,* 1975, 290–297.
69. MURAVCSIK, M.J. *and* MURUGESAN, P. Some results on the function and quality of citations. *Social Studies of Science, 5,* 1975, 86–92.
70. WEINSTOCK, M. Citation indexes. *In:* KENT, A. *and* LANCOUR, H., *eds. Encyclopedia of library and information science.* Volume 5. New York: Dekker, 1971.
71. PRABHA, C.G. Some aspects of citation behaviour: a pilot study in business administration. *Journal of the American Society for Information Science, 34,* 1983, 202–206.
72. CLEVERDON, C.W., MILLS, J. *and* KEEN, E.M. *Factors determining the performance of indexing systems. Volume 1: Design.* Cranfield: College of Aeronautics, 1966.
73. CLEVERDON, C.W. *and* KEEN, E.M. *Factors determining the performance of indexing systems. Volume 2: Test results.* Cranfield: College of Aeronautics, 1966.
74. MOREHEAD, D.R. *and* ROUSE, W.B. Models of human information seeking. *Information Processing and Management, 18,* 1982, 93–205.
75. MOREHEAD, D.R. *and* ROUSE, W.B. Human computer interaction in information seeking tasks. *Information Processing and Management, 19,* 1983, 243–253.
76. ROUSE, W.B. *and* ROUSE, S.H. Human information seeking and design of information systems. *Information Processing and Management, 20,* 1984, 129–138.
77. ROUSE, W.B., ROUSE, S.H. *and* MOREHEAD, D.R. Human information seeking: online searching of bibliographic citation networks. *Information Processing and Management, 18,* 1982, 141–149.
78. APTED, S.M. General purposive browsing. *Library Association Record, 73,* 1971, 228–230.
79. AYRIS, P. *The stimulation of creativity: a review of the literature concerning the concept of browsing, 1970–1985.* Sheffield: CRUS, University of Sheffield, 1985. (Department of Information Studies Consultancy and Research Unit (CRUS) working paper no. 5)
80. CELORIA, F. The archaeology of serendip. *Library Association Record, 70,* 1968, 251–253.
81. HERNER, S. Browsing. *In:* KENT, A. *and* LANCOUR, H., *eds. Encyclopedia of library and information science.* Volume 3, New York: Dekker, 1970.
82. LEVINE, M.M. An essay on browsing. *RQ, 9,* 1969, 35–36, 93.
83. HILDRETH, C.R. Online catalogues and public libraries. *Public Libraries, 23,* 1969, 59–60.
84. SCHUMAN, B.A. Interactive amiability: Dialog, Orbit, and BRS under scrutiny. *In:* VONDRAN, R.F., *ed. Productivity in the information age: Proceedings of the 46th ASIS Annual Meeting, Washington, D.C.* White Plains, New York: Knowledge Industry Publications, 1983.
85. HILLS, P.J., HULL, J. *and* PULLINGER, D. *An experiment on the redesign of journal articles for online viewing.* London: BLRDD, 1983.
86. FREI, H.P. *and* JAUSLIN, J.F. Graphical presentation of information and services: a user oriented interface. *Information Technology Research and Development, 2,* 1983, 23–42.
87. PALAY, A.J. *and* FOX, M.S. Browsing through databases. *In:* ODDY, R.N., ROBERTSON, S.E., VAN RIJSBERGEN, C.J., *and* WILLIAMS, P.W., *eds. Information retrieval research.* London: Butterworth, 1981, 310–324.
88. BERTRAND-GASTALDY, S. *and* DAVIDSON, C.H. Improved design of graphic displays in thesauri – through technology and ergonomics. *Journal of Documentation, 42,* 1986, 225–251.
89. SANDERSON, N. *Graphics in information retrieval.* MSc Dissertation, University of Sheffield, 1984.
90. BOVEY, J.D. *and* BROWN, P.J. Interactive document display and its use in information retrieval. *Journal of Documentation, 43,* 1987, 125–137.
91. CURTIS, P.M.J., PICKERING, H. *and* WHITE, M. ALS update. *Vine, 39,* 1981, 34–36.
92. DOVER, M. *Public information in rural areas: technology experiment (PIRATE).* Boston Spa: British Library, 1988. (Library and Information Report 64)
93. NOERR, P.L. *Information navigation. In:* MACCAFFERTY, M. *and* GRAY, K. *eds. The analysis of meaning: Informatics 5.* London: Aslib, 1979, 202–205.
94. NOERR, P.L. *and* NOERR, K.T.B. Browse and navigate: an advance in database access methods. *Information Processing and Management, 21,* 1985, 205–213.
95. BOLL, J. *The input and output of 22 psychological periodicals: a study of bibliographic coverage.* Urbana, Illinois: University of Illinois, 1952.
96. GEROULD, A.C. *and* WARMAN, H.J. Most cited periodicals in geography. *Professional Geographer, 6,* 1954, 6–12.
97. XHIGNESSE, L.V. *and* OSGOOD, C.E. Bibliographical citation characteristics of the psychological journal network in 1950 and 1960. *American Psychologist, 22,* 1967, 778–791.
98. JAKOBOVITS, L.A. *and* OSGOOD, C.E. Connotations of twenty psychological journals to their professional readers. *American Psychologist, 22,* 1967, 792–800.
99. LIN, N. *and* NELSON, C.E. Bibliographic reference patterns in some sociological journals 1965–1966. *American Sociologist, 4,* 1969, 47–50.
100. KIRSCH, I. Psychology's first paradigm. *Journal of the History of the Behavioral Sciences, 13,* 1977, 317–325.
101. KRANTZ, D.L. Research activity in 'normal' and 'anomalous' areas. *Journal of the History of the Behavioral Sciences, 1,* 1965, 39–42.
102. KRANTZ, D.L. Schools and systems: the mutual isolation of operant and non-operant psychology as a case study. *Journal of the History of the Behavioral Sciences, 8,* 1972, 86–102.
103. RUSSETT, B.M. Methodological and theoretical schools in international relations. *American Academy of Political Science Monographs, 10,* 1970, 87–105.
104. SAMELSON, F. From 'race psychology' to 'studies in prejudice': some observations on the thematic reversal in social psychology. *Journal of the History of the Behavioral Sciences, 14,* 1978, 265–278.
105. WEIMAR, W.B. *and* PALERMO, D.S. Paradigms and normal science in psychology. *Science Studies, 3,* 1973, 211–244.
106. PRICE, D.J. DE SOLLA. *Science since Babylon.* Yale: Yale University Press, 1961.
107. PRICE, D.J. DE SOLLA. Collaboration in an invisible college. *American Psychologist, 21,* 1966, 1011–1017.
108. ALLEN, T.J. *Sources of ideas and their effectiveness in parallel R and D projects.* Cambridge, Massachusetts: MIT, 1965. (Alfred P. Sloan School of Management: Report no. 130–65)
109. ALLEN, T.J. *Managing the flow of technology: technology transfer and the dissemination of technological information within the R and D Organisation.* Cambridge, Massachusetts: MIT, 1977.
110. ALLEN, T.J. *and* COHEN, S.I. Information flow in research and development laboratories. *Administrative Science Quarterly, 14,* 1969, 12–19.
111. GARFIELD, E. *Current Contents* search: new online version of *Current Contents* expands your coverage and retrieval options. *Current Contents, 50,* 1987, 3–9.
112. CONKLIN, J. Hypertext – an introduction and survey. *Computer, 20,* 1987, 17–41.
113. HYATT, H. *Hypertext: a review and evaluation.* MSc Dissertation, University of

Sheffield, 1988.
114. HJERPPE, R. Project Hypercatalog: visions and preliminary conceptions of an extended and enhanced catalog. *In:* BROOKES, B.C., *ed. Intelligent information systems for the information society: proceedings of the Sixth International Research Forum in Information Science (IRFIS 6), Frascati, Italy, 1985.* Amsterdam: Elsevier, 1986.
115. MOON, C. Computerized personal information systems for research scientists. *International Journal of Information Management, 8,* 1988, 265–273.
116. BELKIN, N.J. *and* VICKERY, A. *Interaction in information systems: a review of research from document retrieval to knowledge based systems.* Boston Spa: British Library, 1985. (Library and Information Research Report no. 35)
117. BENNETT, J.L. The user interface in interactive systems. *Annual Review of Information Science and Technology, 7,* 1972, 159–196.
118. BORGMAN, C.L. Psychological research in human computer interaction. *Annual Review of Information Science and Technology, 19,* 1984, 33–64.
119. BRENNER, L.P., HUSTON-MYAMOTO, M., SELF, D.A., SELF, P.C. *and* SMITH, L.C. User computer interface design for information systems: a review. *Library Research, 2,* 1981, 63–73.
120. MAGUIRE, M. An evaluation of published recommendations on the design of man-computer dialogues. *International Journal of Man Machine Studies, 16,* 1982, 236–261.
121. MARTIN, T. Information retrieval. *In:* SMITH, H.T. *and* GREEN, T.R.G., *eds. Human interaction with computers.* London: Academic Press, 1980, 161–175.
122. MORAN, T.P. Guest editors introduction: an applied psychology of the user. *ACM Computing Surveys, 13,* 1981, 1–11.
123. RAMSEY, H.R. *and* GRIMES, J.D. Human factors in interactive computer dialogue. *Annual Review of Information Science and Technology, 18,* 1983, 29–59.
124. THOMPSON, D.A. Interface design for an interactive retrieval system: a literature survey and a research description. *Journal of the American Society for Information Science, 22,* 1971, 363–373.
125. KATTER, R.V. *On-line user of remote access citation retrieval services.* (Report TM(L)). Santa Monica, California: System Development Corporation, 1970.
126. SUMMITT, R.K. Dialog and the user: an evaluation of the user interface with a major online retrieval system. *In:* WALKER, D., *ed. Interactive bibliographic search: the user/computer interface.* Montvale, New Jersey: AFIPS Press, 1971.
127. TIMBIE, M. *and* COOMBS, D.H. *An interactive information retrieval system: case studies on the use of Dialog to search the ERIC document file.* Stanford: Stanford University, 1969.
128. MARCUS, R.S., BENEFELD, A.R. *and* KUGEL, P. The user interface for the Intrex retrieval system. *In:* WALKER, D., *ed. Interactive bibliographic search: the user/computer interface.* Montvale, New Jersey: AFIPS Press, 1971.
129. MARTIN, T.H. *and* PARKER, E.B. Designing for user acceptance of an interactive bibliographic search facility. *In:* WALKER, D., *ed. Interactive bibliographic search: the user/computer interface.* Montvale, New Jersey: AFIPS Press, 1971.
130. MARTIN, T.H., CARLISLE, J.H. *and* TREU, S. The user interface for interactive bibliographic searching: an analysis of the attitudes of nineteen information scientists. *Journal of the American Society for Information Science, 24,* 1973, 142–147.
131. MCALLISTER, C. *and* BELL, J.M. Human factors in the design of an interactive library system. *Journal of the American Society for Information Science, 22,* 1971, 96–104.
132. WALKER, D.E., *ed. Interactive bibliographic search: the user/computer interface.* Montvale, New Jersey: AFIPS Press, 1971.

133. SMITH, K. *An investigation of the information seeking behaviour of academics active in the field of English literature.* MA Dissertation, University of Sheffield, 1988.

APPENDICES

Appendix 1: Research/teaching interests of the social scientists interviewed

MRC/SSRC Social and Applied Psychology Unit

1A Personnel psychology; social psychology of health and illness; unemployment; entry into work; youth training
2D Job design; new technology; organisational behaviour
3B Psychological effects of unemployment
4A Psycholinguistics; human factors in programming and office computing
5D Effects of introduction of new technology in manufacturing on job satisfaction and mental health
6B Relationship between unemployment and health and psychological well being
7A Life span development; work role transitions; organisational behaviour; employee relations
8A Organisational psychology; occupational stress; unemployment; stress in and out of work
9B Human computer interaction; psycholinguistics; cognitive ergonomics
10A Clinical psychology; psychotherapy process and outcome; stress in the working population
11A Computer controlled decision making processes; man-machine communication; computers and decision making
12A Industrial psychology; employee participation; social and occupational psychology; job design; new technology

Psychology

1A Cognitive/perceptual and motor development in children
2A The development of written language and the acquisition of communicative skills; language and communication and communicative competence in normal and handicapped children
3A Computer models of visual function; binocular vision
4A Human memory development and learning; cognition; computer aided learning
5A Neurobiology of motivational systems; psychopharmacology
6A Behavioural pharmacology; neurochemical correlates of behaviour; behavioural genetics; psychobiological aspects of occupational stress
7A Quantitative methods; laterality
8A Early social development; play; human ethology and sociobiology
9A Environmental psychology; social psychology of drug dependence

Education

1D Teaching of English; teaching of drama and its role in education; curriculum studies; media and creative studies

2D Philosophy of education; philosophy of science; philosophy of the curriculum; relationships between philosophy and the social sciences; philosophy of language
3C Sociology of education; educational policy; school and teacher effectiveness; evaluation of educational institutions; research methods
4D Child development; primary education, especially early years and preschool; parental involvement in the teaching of reading; psychology of thinking
5D Science education in primary and secondary schools; curriculum development and evaluation; individualised instruction; learning of scientific concepts; research methods
6D Psychology of education; classroom interaction and classroom processes; adolescence; pastoral care and counselling; research methods
7C Social psychology of education; special education; pastoral care and counselling; qualitative research methods
8D Educational psychology; adult learners; guidance and counselling; special education

Continuing Education
9D Industrial relations; 19th and 20th century labour history; educational provision for people in work
10D Local government and urban politics, especially public participation and community education
11C Local history within South Yorkshire and North Derbyshire

Economics
1C Consumer behaviour; family aspects of buying behaviour; marketing
2D Wage determination; prices and incomes policy in control of inflation; the long wave or Kondratief cycle
3D Accounting theory; management information systems

Economic and Social History
4C History of economic thought; economic history of South Africa and Eastern Europe
5C History of economic policy; economic theory and economic policy in Britain

Geography
6C Agricultural and historical geography of the Mediterranean area; history of cartography; the Venetian cartography of Crete
7C Precipitation origins and analyses; micro-climatology; wind chill effects in upland areas; cave climates

Politics
8C Russian and East European history and politics, especially medieval and 20th century

9C Political economy; Marx and Marxism; the Conservative Party and British economic policy; development of political studies in Britain

Sociology
10C Community care; relationship between informal and statutory care
11D Social planning; research, policy and planning in the personal social services; old age and social policy; the role of research in policy making
12C Sociology of the police; deviance; religion
13C Client studies in social services and child care; social work skills and methods; computing and video use in social work
14C Sociological theory; social and applied philosophy; sociology of sport, leisure and health; moral and physical education

Prehistory and Archaeology
15D Physical and chemical properties of soil as a source of evidence for ancient land use, especially in respect of ancient field systems and settlements

Appendix 2: interview guide

Research/teaching interests
What are your principal research and teaching interests?
How long have you been working on these topics or in this area?
How did you commence work on these topics?
How do you keep up to date with developments relating to this topic?
What criteria do you employ when assessing whether to follow up material?
How do you keep up to date with other developments in the field?
How would you approach the task of moving on to a new topic but in a closely related area?
How would you approach the task of inducting a research assistant or research student into the area?
Could you identify key ideas, authors, to send a research assistant or research student to?
How would you approach the task of moving on to a topic in an area about which you knew nothing?

Characteristics of information use
What are the main sources of information for your work?
Are there any sources which are of particular importance?
Are there any distinctions between the sources or the material which are of particular importance to you?
Which is the most important type of information source: books, journals, reports, conference proceedings, newspapers, etc.
Which are the principal ways you have employed, or intend to employ to publish your own results?
If it is intended to publish the results in journals are these the same ones as those followed?