The Derivation of a Behavioural Model for Information Retrieval System Design

by

David Ellis

Department of Information Studies

Thesis submitted for the Degree
of Doctor of Philosophy at the
University of Sheffield

August 1987

ABSTRACT

This study outlines the derivation of a behavioural model intended to underpin thinking on questions concerning the design of information retrieval systems for academic social scientists. The historical background of information retrieval research is reviewed and the behavioural assumptions made in that research, characterised as the information retrieval model, are examined. Developments in the research tradition following from the early empirical tests on information retrieval systems are discussed, and problems experienced in the attempt to develop a coherent and practically useful research programme around the notion of relevance as a quantitative concept are analysed.

A major alternative to work employing the information retrieval model has been the cognitive approach to information retrieval system design. In research so far undertaken in this approach the system has had to construct a cognitive model of the searcher's requirements in order for retrieval to take place. It is argued that the characteristics of social science information present particular difficulties for building such a model, and that the step of the system building a cognitive model of the searcher's requirements as a pre-requisite for retrieval is unnecessary.

Instead, a behavioural approach to system design is recommended. The information seeking patterns of a variety of academic social scientists, derived from transcripts of interviews, were analysed and broken down into six characteristics, starting, chaining, browsing, differentiating, monitoring, and extracting. These characteristics seemed sufficient to exhaust the different generic features of the various patterns, and to provide a flexible behavioural model to underpin thinking about information retrieval system design.

The features of an exploratory retrieval system for social scientists based on the characteristics of their information seeking patterns are outlined, the extent to which such features are available in existing systems, and the requirements for implementing the model on an operational system are considered. The argument of the study is that the behavioural approach, and the analysis of information seeking patterns into their characteristics, represents a coherent and practically useful basis for the design of exploratory information retrieval systems for academic social scientists.

The Derivation of a Behavioural Model for Information Retrieval System Design

'Perhaps the most important and least considered factor in the design of information storage and retrieval systems is the user of such systems. Regardless of what other parameters are considered in the development of a storage and retrieval mechanism, it is necessary to consider its potential use and mode of use by the persons or groups for whom it is intended; it is necessary either to fashion the system to suit the user's needs, habits, and preferences or to fashion the user to meet the needs, habits and preferences of the system. Both approaches are possible, but the second one, involving education and re-education of the user, is evolutionary and futuristic. A system designed for now should at least be able to serve the present user' (Comment attributed to Saul Herner, 1959: in Allen, 1977:7)

Contents

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 cooperate 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.

Preface

A Note on Method

This study can be conceived, in part, as a test bed for a qualitative, empirically based, and behavioural approach to questions connected with the design and evaluation of information retrieval systems. This is in contrast to the traditional quantitative and probabilistic approaches to such questions. It is also in contrast to more recent work on the development of user models for information retrieval system design in that it is behaviourally not cognitively, and empirically not intuitively based. The behavioural model presented has been derived from analysis of the information seeking activities of the target group not from an a priori set of assumptions or intuitions concerning those activities.

Part of the reason for adopting this approach to the problem was the experience of a previous and much

smaller study which had followed very closely the type of approach to evaluation implicit in the information retrieval model (Ellis, 1982). The characterisation of that approach and some attempt to describe its role in foreclosing explanations of the real problems experienced in the use of retrieval systems has been made elsewhere (Ellis, 1984a: Ellis, 1984b).

This fostered a conviction that to attempt a conventional evaluation of the factors affecting the retrieval effectiveness of computer based information retrieval systems would add little to the established account - and would be likely to be so hedged with reservations as not to be worthwhile undertaking. To have carried out a set of model searches for social scientists on available computerised databases would have been too artificial, but to employ real queries with real relevance judgements - given the experience of the previous study - seemed a method that was fraught with difficulties and likely to lead to little in the way of conclusive results.

The alternative approach seemed to be to turn the problem around. Instead of starting from the information retrieval system and the success or otherwise of social scientists use of it, to start with the social scientists and attempt to build up a picture of the kinds of ways in which they sought information and how the activities that *might be characterised as* 'information seeking' or 'information gathering' were

integrated with their other activities. The intention was to construct a more accurate model of the information seeking activities of the social scientists, and from this to derive recommendations for the design of improved information retrieval systems.

In this respect, although this study is related to other work on the development of user models for information retrieval system design - such as that reviewed by Belkin and Vickery (1985) and by Daniels (1986) - it differs from most of that work in being empirically based on analysis of actual information seeking patterns, and in adopting a behavioural rather than cognitive approach to the problem.

The study owes an enormous debt in terms of method to three American works - Glaser and Strauss's (1967) The Discovery of Grounded Theory: Strategies for Qualitative Research, Diesing's (1971) Patterns of Discovery in the Social Sciences, and Patton's (1980) Qualitative Evaluation Methods. Glaser and Strauss's influence is particularly apparent in the choice of samples for interviewing and in the form of the analysis of the interview transcripts. Diesing's in clarifying the relationship between formalist and empiricist experimentation, and in providing an understanding of the characteristics of holistic theories. Patton's influence came in the actual conduct of the interviews, the adoption of the interview guide approach, and finally in the decision to attempt to

make full transcripts for the purposes of analysis.

The Choice of the Sample

There are numerous studies of the characteristics of social science information and the information seeking activities of social scientists. This literature has been extensively reviewed (Adam, 1971: 1982: Brittain, 1970: 1979: Ellis, 1986: Hogeweg-De-Hart, 1983: 1984: Roberts, 1981). Many of the studies have focussed on the information seeking activities of academic social scientists. Academic psychologists have been studied as part of the large scale American study of information communication and exchange in psychology (American Psychological Association, 1963: 1965: 1969: Garvey and Griffith, 1972: Garvey, et al., 1972: 1984). Associated studies have compared the information and communication

activities of physical and social scientists (Garvey, et al., 1970: 1971: 1972).

Two substantial research programmes undertaken in the United Kingdom also concentrated primarily on the information requirements and provision of information services for academic social scientists. The first of these, the investigation of the Information Requirements of the Social Sciences (INFROSS) project, was mainly concerned with the information requirements of academic social scientists. The second, the Design of Information Systems in the Social Sciences Programme (DISISS) examined issues relating to the provision of secondary services in the social sciences. Summaries of the findings of both these major studies are available (Line, 1971: Bath University, 1980).

Also of note is a research project undertaken by a team under Professor Don Swift on index language design for a social science subject (Swift, et al., 1974). This project led to the development of a novel approach to indexing, the multi-modal approach (Swift, et al., 1977a: 1977b), which was then applied, although apparently later abandoned, in the construction of the index for Sociology of Education Abstracts. The project also gave rise to a brief debate on the relationship between sociology and information science (Swift, et al., 1978a: 1978b: 1979: Watson, et al., 1973). Of related interest, are accounts of the perceptions and information seeking activities of scientists and

-echnologists (Allen, 1977: Martyn, 1986: Meadows, 1987: Rowland, 1982), and of humanities scholars (Stone, 1982). A particularly interesting study, in this respect, is that by Skelton, (1973) which compared the results of science user studies with the INFROSS findings.

However, despite the useful background that these studies provide for an understanding of the characteristics of social science information and social science information use, both generally and with particular reference to academic social science, none of the studies provide the kind of detailed account of the perceptions of academic social scientists of their information seeking activities, from their point of view, and seen as a whole, necessary for this study. It was this type of micro-level information about the activities and perceptions of the academic social scientists which was required for the analysis of the system requirements. In this respect, the detailed individual studies from which the analysis of the system requirements are derived, represent a unique resource for the study of information use by academic social scientists.

The decision to focus on academic social scientists was partly dictated by interest, partly by the availability of a good sample close at hand, and partly by the desire not to become involved in the added complexity, and different problems, of

information use of practitioners and policy makers. The restriction to academic social scientists working in a university as opposed to a polytechnic or college of higher education environment was based on the fact that the study was not intended to be offering some kind of comparative study of academics in different institutional settings. Therefore, it seemed valid to focus on social scientists working in a university environment - who in the event provided more than enough variation.

The choice of Sheffield was decided on the grounds of convenience and ease of access. However, the University does represent a fairly typical 'redbrick' environment with a broad span of subject interests and mix of undergraduate and postgraduate teaching. While the existence of differences between the information seeking activities of social scientists in a university as opposed to a polytechnic or college of higher education environment cannot be ruled out, the existence of such a difference between social scientists in Sheffield and those in other universities seems remote enough to discount.

The choice of the first group of social scientists to interview, however, was affected 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 cooperate 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 simply too busy for a convenient interview to be arranged. The choice of both this and the other samples for interview was based on Glaser and Strauss's technique of theoretical sampling. It became clear, for example, that the initial choice of social scientists 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 45 minutes to 2 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 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 some a priori notion that it might 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. This comprised four from Sociological Studies - two from Sociology and two from Social Administration - two from Economic and Social History, two from the Division of Economic Studies, two from Geography, two from Political Theory and Institutions, two from Education, and two from 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. They were on the whole a little shorter - lasting between 45 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 with the help of the University Social Sciences Librarian David Jones. The 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 cooperation of David Jones 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, this group included one student who had made a transition from advanced academic study in prehistory and archaeology to take the MSc Information Studies (Social Sciences) course. Brief outlines of the subject interests of the social scientists interviewed are given in appendix one.

For clarity of exposition the responses of the individual social scientists interviewed have been treated together in broad subject groupings, individuals are then identified by number within those groupings, however, the individual responses have been coded so that it is possible to identify from which sample set the individual response was taken. The codes relate to the source used to identify the original sample sets.

A MRC/SSRC SAPU and Psychology staff list;

B MRC/SSRC SAPU research assistants;

C Social Science Faculty staff list;

D Library computer search records.

Interviewing and Analysis

Although there has been extensive study of information seeking by academic social scientists there seemed little to go on in terms of guidance for designing a schedule for such a group that would not lead to serious distortion of the response. Informal, semi-structured, in depth interviews seemed the best means of obtaining the right kind and level of information with the least pre-judgement of the results. Ultimately the decision was made to employ an extremely simple format (appendix two), the questionnaire being used as a guide to ensure that no major part of the social scientists' information seeking activities was missed. The intention was not to set up a situation where the individual social scientists ended up talking rather artificially about their 'information behaviour' or some other such abstraction, or to talk, in isolation, about their use of the library. The intention was for them to talk about their work and the sorts of things they did which might be understood as having an information component.

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 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 (1967), with further practical guidance on carrying out the analysis being derived from Glaser (1978) and Turner (1981). 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. As the analysis progressed it became clear that certain categories were more or less synonymous with others and that other candidate categories were better treated as properties of broader categories. In the end six major categories seemed to satisfactorily subsume the important characteristics of the information seeking patterns.

1. Starting
2. Chaining
3. Browsing
4. Differentiating
5. Monitoring
6. Extracting

It should be noted, that although Glaser and Strauss's approach has served as a guide to the carrying out of this study there are differences. For example, it was not considered feasible to carry out triangulation of methods, partly because of the practical difficulties involved, but also because it simply did not seem necessary - the interviews and the transcripts seemed to provide an authentic picture, in themselves, of the activities of the social scientists, and a sufficiently accurate and detailed picture of the information component of those activities could be distilled from those accounts.

The most important difference, however, is that the result is not in the full sense a grounded theory of the information seeking activities of the social scientists. The theoretical component to this study emerges from from the general postulates that

1. it is exploratory capability which provides the soundest base for designing information retrieval systems which will mesh with the information seeking patterns of social scientists:

2. and that it is the behavioural rather than the cognitive aspects of those information seeking patterns which serves as the better guide for the design of such systems.

The contribution of the concept of grounded theory comes in the derivation of the characteristics of the behavioural model which is advanced, in that the characteristics are empirically derived from the analysis of the interviews.

Finally, throughout this study the term 'information retrieval systems' is used to describe the, predominantly, documentary retrieval systems which have been the subject of most work in information retrieval research. This usage has been criticised as confusing by Martin (1980) who distinguishes between 'table lookup', 'document retrieval', 'decision support' and 'question answering' systems. Nevertheless, the application of the term 'information retrieval' to systems which are primarilly concerned with document retrieval corresponds to the historical employment of the term in the field, and, even with the recent interest in the application of expert system techniques in information retrieval research, documentary retrieval systems have continued to be the

main focus of research interest, therefore, the convention of applying the term 'information retrieval research' to such work has been followed here.

Chapter One

Laboratory Testing of Information Retrieval Systems

The Historical Background of Information Retrieval Research

The origins of information retrieval research can be traced back to 1953 when separate groups of tests were carried out in Britain and the United States

evaluating the performance of the, then controversial, 'uniterm' system of Mortimer Taube against more conventional approaches to subject indexing and retrieval. These two tests were the Cranfield - Uniterm test undertaken in the United Kingdom and described by Thorne (1955), and the Armed Services Technical Information Agency (ASTIA) - Uniterm test carried out in the United States which was reported by Gull (1956).

However, 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 testing and research as a sub-discipline of library or information studies. Whether the results of the Cranfield tests were accepted or rejected subsequent work on the question of the effectiveness of information retrieval systems had to take the Cranfield tests, or the approach to testing retrieval effectiveness derived from those tests, into account.

The tests themselves have been fully described (Aitchison and Cleverdon, 1963: Cleverdon, 1960: 1962: Cleverdon, et al., 1966: Cleverdon and Keen, 1966), summarised (Cleverdon, 1967: 1971: Sparck Jones, 1981a: 1981b: Lancaster and Mills, 1964), and criticised (Harter, 1971: Mote, 1963: Rees, 1965: 1967: Richmond, 1963: Sharp, 1964: 1967: Swanson, 1965: 1971: Vickery, 1966: 1967). They established the principle that arguments about the relative merits of different

retrieval system designs had to be empirically grounded. In this respect, they mark an historical change in consciousness from a philosophical and speculative approach to information systems design to and empirical and experimental one.

The tests also provided the theoretical framework within which the sub-discipline of information retrieval research subsequently developed. This has had unfortunate consequences for the progress of the sub-discipline. For, underlying the methodology of the tests was a set of behavioural assumptions, which may, with the benefit of hindsight, be seen to have both restricted the scope of the sub-discipline and hindered progress in it. However, before an attempt is made to examine those behavioural assumptions, some account of how the sub-discipline developed may help to illuminate the relationship of those assumptions to the later, as well as to the earlier work.

The Test Collection Method

The Cranfield tests were empirical tests in which an effort was made to hold certain factors constant while others were varied and controlled observations carried out, in this sense they can be said to have been empirically driven. Later work in the same tradition has been far more explicitly theory driven, and the impetus to develop an information retrieval theory has been strong (Robertson, 1977). Much of the work carried out has employed computational methods to manipulate data previously gathered from laboratory tests. The results of previous laboratory tests, converted into machine readable form and referred to as test collections, being used as the basic data set for the application of the computational techniques. There are several such collections in regular use (Griffiths, et al., 1986), and the results of the the second series of Cranfield tests have been frequently employed in this form.

It should be noted, that a test collection, in this context, is not merely a collection of documents. Typically it consists of three elements -

1. a set of document representations or sources of index descriptions of documents;

2. a set of queries;

3 a set of relevance judgements.

It is not intended to go into detail on the problems involved in setting up such test collections - which have been described by Sparck Jones and Van Rijsbergen (1976) - but rather to consider the questions of the relationship of this research method to the tradition of laboratory testing from which it developed, and how far use of the method affects the underlying theoretical and behavioural assumptions of the earlier tests.

The explicitly theoretical approach to information retrieval research, and the widespread use of sophisticated computational techniques and mathematical models, made possible by the employment of the test collection method permitted a formalism in approach that contrasts with the strong empiricism of the Cranfield tests. For example, in the Cranfield tests the inverse relationship between precision and recall is an empirically derived hypothesis, whereas, as Van Rijsbergen (1981) notes

> 'under some mathematical models in information retrieval the trade-off is a necessary one, and not simply observed empirically' (Van Rijsbergen, 1981: 33)

Drawing a clear distinction between formalist and empiricist employment of mathematical models in information retrieval research is difficult, as with other work with mathematical models in the social sciences the demarcation between the two kinds of experimation can be

> 'more a matter of difference of interpretation and of ultimate aim than of substance, though differences of substance do remain' (Diesing, 1971:85)

Nevertheless, whether the intention behind work with mathematical models in information retrieval research is empiricist or formalist in orientation, the control over the variables and the sophistication of the computational techniques employed allows theories and hypotheses to be developed at a level of abstraction that can lead to unease as to the applicability of the results to operational systems.

Oddy (1981) had certainly come to the view that there was serious risk that the approach had become too far removed from operational reality, Remarking that

> 'Some researchers appear to believe that their understanding of the information retrieval problem through mathematical theories has attained a plateau - quite a lofty plateau! In this work there has been too little reference to the real world to justify this optimism' (Oddy, 1981:175).

and felt,

> 'uneasy about solutions to the realism problem which involve larger and better planned test collections and repetition of tests on a number of different test collections' (Oddy, 1981:170))

Heine too considered there were dangers inherent in this shift towards experimentation with formalist mathematical models

> 'it needs to be clearly understood that simulation work, as a species of theoretical work, is always dependent and sometimes critically dependent on valid data from laboratory or operational system studies. Without a continuous stream of such data, reacting symbiotically with simulated theoretical study both will be poorer; we will have a theory of unproven applicability, and operational systems of unproven optimality' (Heine, 1981:197).

The important thing would seem to be to establish whether the results of the earlier laboratory tests or employing the test collection method apply in operational environments, and this has proved notoriously difficult (Cleverdon, 1972: Lancaster, 1968).

Nevertheless, despite the differences in approach and emphasis between early laboratory tests, and later work employing the test collection method, the manner in which the test collection method has been applied and developed has not led to any questioning of the behavioural assumptions made in the early laboratory tests. The behavioural assumptions underlying research

employing the test collection method are identical to those of the laboratory tests which provide the data set. Because of their implicit role in, and continuing significance for, this tradition of information retrieval research, the behavioural assumptions made may be referred to as constituting the information retrieval model (Ellis, 1984a: Ellis, 1984b).

The Information Retrieval Model

The information retrieval model is the behavioural model embedded in the testing procedure of the Cranfield and following tests. This behavioural model was implicit in the method of the tests, but was not less significant for not having been made explicit. The fact that it was built into the testing procedure has perhaps led to it being both more pervasive and less open to recognition and criticism than if it had been

made explicit.

In the model a user recognises an information need. The user then comes to an information retrieval system with a request based on that need. The retrieval system matches the request against representations of documents in the system. The task of the system is seen as that of presenting to the user the text or texts most likely to satisfy his need. The user examines the representations of the texts presented and judges them for relevance. The intention is that some or all of the documents will either partially or wholly satisfy the user's information need.

The contribution of the different elements of the model to the eventual retrieval outcome is altered in work employing relevance feedback, in that relevance judgement is employed to modify the original query formulation (Robertson, et al., 1986: Robertson and Sparck Jones, 1976: Sparck Jones, 1979a: Sparck Jones, 1979b), but this work does not affect the basic configuration of the model. The assumptions of this model have been given explicit formulation by Robertson and Belkin (1978) and by Belkin (1981). The characterisation of the model as the information retrieval model, and its role in information retrieval research has been examined by Ellis (1984a: 1984b). It should be underlined that the relationship of the information retrieval model to the assumptions made in the Cranfield tests is a direct one. The principal

features of the model are identical to those of the testing procedure. That is, they are not grounded in empirical observation, but represent an abstraction from, or idealisation of, the situation implicit in that procedure. So that, the information retrieval model represents an explicit formulation of the implicit behavioural assumptions made in the tests.

This model has had considerable direct and indirect influence on information retrieval research. Its direct influence came from the employment of the Cranfield results in the form of a test collection. Its indirect influence derives from the fact that the model of the information retrieval situation assumed in the Cranfield tests was implicitly accepted in later work, both in subsequent empirical studies and in the employment of data from those studies in the form of test collections.

The behavioural assumptions of the information retrieval model have informed a whole tradition of research into the design and testing of information retrieval systems. The majority of the research carried out in the field of information retrieval research following the Cranfield tests did not question the assumptions of the information retrieval model, even where those assumptions were recognised. To a large extent this has led many researchers to focus on problems posed by the information retrieval model, rather than examining the assumptions of that model, or

considering whether alternative approaches to the question of information seeking behaviour, than that embodied in the information retrieval model, might suggest different solutions to the problem of how to increase the effectiveness of information retrieval systems.

The different components of the model have been the subject of many studies. The concept of 'information need' has been reviewed by a number of writers (Allen, 1969: Belkin and Vickery: 1985: Brittain, 1970: Lin and Garvey, 1972: Martyn, 1974: Paisley, 1968: Wilson, 1981). Other writers have examined questions concerning the derivation of request statements and their significance for information retrieval research (Belkin and Vickery, 1985: Heine, 1980: Saracevic, 1978: Taylor, 1968). However, it is the concept of relevance which has been the subject of most work and comment, and it is proposed here to concentrate on this aspect of the information retrieval model, in order to demonstrate that distortion of behavioural reality is a necessary, not an accidental, feature of the form of testing procedure derived from Cranfield.

The literature on relevance is extensive and has been reviewed by Saracevic (1970a: 1970b: 1975). This literature is often of intrinsic interest, and much of it is related to questions concerning its employment in evaluation and testing. However, in the present

context, this literature will only be referred to where it connects with specific issues raised by its employment in the tradition of research which has followed from the Cranfield tests. The concept of relevance has been significant to this research in two ways -

1. in its employment as a performance criterion;
2. and its use in relevance feedback systems.

These two uses raise some connected questions, but before exploring these questions it is perhaps best to see how its use in the two roles differs.

Relevance as a Performance Criterion

The first use of relevance as a performance criterion for retrieval system evaluation was in the ASTIA - Uniterm test (Gull, 1956). Two groups of testers, one consisting of indexing staff of ASTIA, the

other staff of Mortimer Taube's company, Documentation Incorporated, separately indexed and then searched the ASTIA collection, which then consisted of around 15,000 documents, with 93 requests sent to ASTIA in the normal course of its activities. The measure of effectiveness was relevance of documents to the question. This appears to be the first appearance of relevance as a performance criterion for retrieval system evaluation, and, therefore, at least in the United States, relevance and information retrieval testing make their entrance simultaneously.

In retrospect this was perhaps a rather inauspicious beginning, as the test apparently broke down in disarray over the question of relevance judgement. The two groups of testers were unable to come to an agreement, with the Documentation Incorporated group holding that Uniterm had performed better, and the ASTIA group maintaining the opposite. It is unfortunate that the form of the dispute between the two groups is overshadowed by the fact that neither of the parties could be said to be entirely disinterested in the outcome of the tests. Nevertheless, the breakdown of the test is significant since they were, whatever their possible different predilections as to the actual outcome, supposed to be employing the same criterion of assessment, and that not only were they unable to to agree on the question of relevance judgement, but also that there was no

decision procedure to which they could appeal to resolve their differences.

Cleverdon noted this problem with the ASTIA - Uniterm test in the report of the first Cranfield tests.

> 'It would appear that when the searches were completed, each organisation looked at the documents which had been retrieved and decided on those which were relevant to each particular question. The two groups then met to compare results. Immediately they came up against the problem of deciding what was relevant and found that they were quite unable to agree on this point. Each group had its own interpretation of the question and therefore its own views as to the relevance of the documents' (Cleverdon, 1962:7)

Adding, that,

> 'It was quite obvious that, whatever other weaknesses the test might have it was essential that it should not get bogged down in the quagmire of arguments concerning relevancy' (Cleverdon, 1962:7)

The Cranfield - Uniterm test (Thorne, 1955) - which preceded the first Cranfield tests proper - had adopted a rather different approach to testing to that followed in the ASTIA - Uniterm test. In the Cranfield - Uniterm test a limited sub-set of 200 documents was searched with 40 artificial questions, and the criterion of effectiveness employed was that of success in retrieving source documents, that is the documents from which the queries employed had been artificially

derived. However, this form of test procedure had two obvious limitations.

1. Only a single figure performance value was derived, based on the retrieval of the source document. No corresponding figure was given for the retrieval of non-source documents, or other 'relevant' source documents.

2. A term-based system might be said to be favoured over a a concept-based system if the queries were influenced by the terms used in the source documents.

It is interesting to note that these limitations were the basis of some of the most serious criticisms made of the first series of tests later carried out at Cranfield, and that the methodology of that series of tests was virtually identical to that employed in the Cranfield - Uniterm test, although on a much larger scale. The one real methodological difference being that the measure of effectiveness was broadened slightly to count retrieval of the source document, or document at least as relevant as the source document.

Cleverdon, when explaining the rationale for the methodology employed in those tests, makes no explicit reference to the earlier Cranfield - Uniterm test, on which it is clearly based, but does refer to the ASTIA - Uniterm test indicating the abyss into which that test had fallen. But as with the earlier Cranfield - Uniterm test the limitations of the procedure adopted

were that again

1. Only a single figure performance value was obtained, no corresponding figure was given for the retrieval of non-source documents.

2. Term based systems might be said to be favoured over concept based systems if the queries are too closely related to the actual terms used in the source documents.

Nor did Cleverdon's attempt to avoid arguments concerning relevance entirely succeed. Through the employment of source documents the internal conduct of the first series of tests was kept free of arguments concerning relevance. But the use of source documents in two roles, to derive the queries and to serve as the basis for measuring effectiveness, was a widely criticised aspect of the first Cranfield tests.

In following tests carried out at, or in collaboration with, the Cranfield Institute, the contribution of relevance judgement to the evaluation of retrieval effectiveness increased. The later Cranfield - Case Western University test (Aitchison and Cleverdon, 1963), which was otherwise methodologically identical to the first series of Cranfield tests, employed retrieval of the source document, or document at least as relevant as the source document, or document less relevant than the source document. By the time of the second major series of tests carried out at Cranfield, retrieval of relevant documents had

supplanted retrieval of source documents as the measure of effectiveness.

The second series of Cranfield tests were rather different in form from the earlier tests, in that they were not of operational systems, instead a number of different types of indexing language with varying terminologies and structures were constructed (Cleverdon, et al., 1966: Cleverdon and Keen, 1966). The tests were designed to investigate the effect that different generic index language features had on retrieval performance. As with the earlier tests carried out at Cranfield, the document base consisted of material on aeronautics, in this case 1400 items. A total of 211 requests were obtained by asking authors of selected published papers (the base documents) to reconstruct the questions which originally gave rise to the writing of these papers. However, in order to speed up the conduct of the tests, most of the tests were carried out employing smaller subsets of documents and queries, the smallest consisting of 200 documents and 42 requests.

The measure of effectiveness was now explicitly relevance based. The testing procedure employed a two stage judgement of relevance. First students of aeronautics searched the entire document collection. The documents they identified as relevant - plus the references given in the base document were then sent for final relevance judgement by the author/requestor.

This return to an explicitly relevance based measure of effectiveness meant that Cleverdon had once again to consider what had gone wrong with its employment in the ASTIA - Uniterm test.

Cleverdon noted that the point at which problems arose was when the two groups had come together

> 'to analyse the output of their searches to find which documents were relevant' (Cleverdon, et al., 1966:9)

> 'Up to this point, everything appears to have gone according to plan. The final stage was intended to be a comparison of the output of the two sets of searches, in order to find which system had been successful in obtaining more relevant documents. The problem which arose at this final stage was that neither group was willing to accept the relevance assessments of the other group; rumour has it that at the end of the second day of discussion, the two groups were still arguing about the meaning of the first search question. No real blame can be fixed on those who organised the test; in 1952 it was not unreasonable to think that two groups of intelligent people would, without serious difficulty, be able to come to an amicable agreement as to which documents were relevant to a particular question' (Cleverdon, et al., 1966:10-11)

Cleverdon gave the question of the form of relevance assessment to be employed in the second series of tests considerable thought. Weighing up the various possibilities he concluded -

> 'The theoretical ideal is...the use of actual

> questions with relevance assessment made at the time by the questioner from the complete texts...For this project it was considered that the nearest to the ideal would be the combination...questions which had been asked, with a relevance assessment being made by the questioner who would be a scientist' (Cleverdon, et al., 1966:15)

Cleverdon drew a a distinction between two different forms of relevance judgement, 'stated relevance' and 'user relevance', according to whether the judgement was based on the searcher's request statement or underlying need.

> 'The enquiry is expressed in the form of a 'stated requirement'...quite often the stated requirement is far removed from the real needs of the questioner...These two types of relevance are called 'user relevance' and 'stated relevance'. The former can only be decided by the questioner himself, but 'stated relevance' can be determined by anybody with reasonable knowledge of the subject field...in this particular project an endeavour was made to simulate 'user relevance' decisions' (Cleverdon and Keen, 1966:256-257)

The notion that ascertaining 'stated relevance' is unproblematic seems to ignore the problems experienced in the ASTIA - Uniterm test, for, as Cleverdon states, problems had arisen over interpretation of the meaning of the first search question. It is also questionable how far the procedure adopted can be said to have simulated 'user relevance'. Certainly, the final relevance judgement may be said to have been based on the questioners attempting to reconstruct the need

underlying their original request, but the student pre-selectors can only have had the request statements on which to base their judgements. This raises the question whether the two groups were doing the same thing, or doing different things in their relevance judgements.

The difficulties involved in this procedure did not go unremarked by the testers themselves, and Cleverdon expressed reservations both about the decision to attempt to simulate 'user relevance' decisions and about the relevance judgement procedure adopted.

> 'Whether the original decision to simulate user relevance decisions was correct has already been considered (Vol. 1, page 114) and tentatively the conclusion was there reached that it might have assisted the interpretation of the test results if, instead, stated relevance decisions had been used. On the whole, this is a view to which we would still subscribe but for one fact. If stated relevance decisions had been used, and assuming that the test results had shown the similar superiority of singe term natural language, then it would have been virtually impossible to refute an argument that the results were unduly influenced by the relevance decisions' (Cleverdon, and Keen, 1966:257)

There is no doubt that the method employed to simulate 'user relevance' judgements was ingenious, even if the distinction between 'stated relevance' and 'user relevance' is problematic, and despite the fact that it is possible to overlook the objective of

simulating 'user relevance' when reviewing the tests, as Belkin appears to,

> 'The Cranfield experiments, and others, recognised and attempted to control for the need problem by eliminating it entirely through the use of artificial questions (that is questions without underlying needs). Then the relevance judgements were carried out in an 'objective' manner untainted by individual differences among the variable users. This strategy is useful in that it explicitly recognised the difficulty of dealing with individual information needs. The problem is that there is no a priori reason to suppose that the performance of a system measured in this way correlates at all well with performance as evaluated by posers of real questions' (Belkin, 1981:55)

However, in the actual conduct of the tests, problems arose both in the missing of relevant documents by the student pre-selectors, and in difficulties expressed by the final assessors themselves in judging the relevance of documents (Cleverdon, et al., 1966). A difficulty, in this respect, which is of particular interest was the comparatively low number of papers cited by the authors which they considered were relevant to the questions, as Cleverdon observed

> 'The relevance assessments that the authors made of their own cited papers reveal some information on the citation habits of authors, but any observation can only be made within the limits of this situation, in which in most cases only a selection of the cited papers was used. A few of the authors assessed all their cited papers as not relevant to the basic questions, and one explicitly stated that he did not find any relevant at all. An analysis of 174 of the

> basic questions, more than was ultimately used, shows that 36% of the cited papers submitted were assessed as not relevant, and if marginally relevant papers are included, the figure is 52%. The results from the basic 118 questions give figures of 28% and 46% respectively. It may be concluded that about half the references in an author's paper are not included in connection with the main problem of the paper' (Cleverdon, et al., 1966:30)

The difficulties experienced in the judgements of relevance have the consequence, pointed out by Cleverdon, that

> 'the operational performance characteristics of the system being tested will almost certainly change depending on the combination of questioner and relevance assessor used' (Cleverdon, et al., 1966:15).

Cleverdon concluded that the one remaining problem for information retrieval research concerned the nature and reliability of relevance judgement, and, with reference to the method of the tests themselves, conceded that on

> 'the matter of obtaining relevance assessments...here some reservations have to be admitted concerning the method adopted. This is not to suggest that there is any experimental evidence of their being any better or more satisfactory technique, but rather to say that the matter of relevance assessment is without any doubt, the most difficult intellectual problem - in fact, one of the very few remaining problems - in the evaluation of information retrieval systems' (Cleverdon, et al., 1966:114)

It is worth underlining why, despite having earlier wished to avoid arguments concerning relevance, relevance judgement is central to the methodology of the later tests, and why its employment is so hedged with reservations. The significant point is that - having failed in the attempt to avoid the use of relevance judgement in the tests - the form of relevance judgement employed had to be objective enough to serve as the basis of the measure of effectiveness, but, at the same time, had to appear sufficiently realistic not to invoke the criticism that the results of the tests were merely an artifact of the type of relevance judgement employed.

In the period following the Cranfield tests a considerable body of research was addressed to questions relating to the variability and consistency of relevance judgement, this work has been reviewed by Saracevic (1970b: 1975). The research was inconclusive, the results of some studies seemed to indicate that a large number of factors affected absolute agreement on relevance judgement (Cuadra, et al., 1967), the results of others that absolute differences in relevance judgement might not affect agreement on the ranking of documents by their relevance (Lesk and Salton, 1968). Differences of emphasis and approach make it difficult to effect a direct comparison of the results. Moreover, it is debatable to what extent the results of the studies provide insight into the question of the

relationship between the sorts of relevance judgments made in the laboratory and the sorts of relevance decisions made in real life, as Lancaster points out

> 'many factors affect the decision of a user as to whether or not a particular item is pertinent to his information needs. These may not be the same factors that would influence his decision were he asked, somewhat artificially, to judge if the same item is relevant to his request statement' (Lancaster, 1979:272)

Before considering this point further, the second major type of use of relevance judgement in information retrieval research will be outlined, that is, its role in providing feedback.

Relevance Feedback

Experimental work evaluating the potential of employing relevance feedback to improve retrieval performance developed from the earlier retrieval system tests. The work is particularly significant in that it circumvents one of the major difficulties inherent in the Cranfield tests - the problem of a user stating his true need in a request. The work has been largely carried out employing test collections (Robertson and Sparck Jones, 1976: Sparck Jones, 1979a: Sparck Jones, 1979b), although more recent work has been carried out employing relevance feedback 'front ends' and using commercial systems (Robertson, et al., 1986). Relevance judgement has a dual role in the work, being employed both to provide the information used to refine the search, and to evaluate the results.

Two tenets underly the research which has been carried out on relevance feedback - that relevance can only be determined probabilistically, and that documents should be ranked in order of their probability of relevance (Robertson, 1977: Robertson and Belkin, 1978: Robertson and Sparck Jones, 1976). Probabilistic approaches to retrieval have been around

a considerable time, as early as 1960 Maron and Kuhns (1960) had advocated a probabilistic approach to the retrieval problem. In the context of work on relevance feedback -

> 'Probabilistic theories assume that relevance is a boolean variable, that is it can take one of two values, denoted: relevant and non-relevant...Other assumptions typically made about relevance are that, for any document-query pair, the relevance judgement is independent of time, and of the other relevance judgements' (Oddy, 1981:169)

In experimental work employing relevance feedback the probability of a document, or group of documents, being relevant, is usually calculated from the presence of terms in documents judged relevant or non-relevant by the searcher. The research has raised a variety of questions concerning how the relevance judgements on which the feedback is based should be made. For example, how much of the document the person should see before making the relevance judgement.

> 'The ideal is clearly the entire text of the document; but again, this is usually out of the question; usually abstracts or titles are used. There has been some work on the prediction of relevance (of full texts) on the basis of titles or abstracts, and it tends to show that titles alone are very bad indicators, abstracts are better but still leave a lot to be desired' (Robertson, 1981:17)

The question of whether the relevance decisions

could be assumed to be independent of each other has also been considered problematic. The difficulty being that even if it were accepted, in principle, that the relevance of an individual document might be connected to the set of documents it was retrieved with, it was not clear how this feature could be successfully operationalised for testing purposes.

> 'If we consider an inferential system answering a question, then several of the basic data units in this system (propositions or whatever) contribute together to the resolution of the question. We are presently working in documentary information retrieval with models of relevance which do not allow for that kind of interaction. But our models of relevance are in some sense clearly inadequate; a possible extension, analogous to the above, would be to allow for a group of texts to be relevant as a group, when individually they are not (or not so) relevant. Such a model would seem to be almost the simplest possible application of semantics to information retrieval; but already it would require us to rethink most of our basic ideas in the design of information retrieval systems' (Robertson, 1979:205).

The assumption that the relevance judgements can be treated as independent of each other entails that the order in which the documents are presented does not affect the judgement of relevance. Robertson considered this assumption was also, in theory, problematic, but noted that no method had yet been employed which might circumvent it.

'The order in which documents are presented

> to the judge may be important. In some sense, it is obviously an over-simplification to regard relevance as something which can be judged for each document independently of the others; one might more reasonably expect the judgement on any one document to be affected by which documents the judge has already seen. Ideally one would try to devise an evaluation method which took this into account; in practice, no such method has yet been used' (Robertson, 1981:17)

A similar difficulty arises with the assumption of consistency. It has to be assumed that the same criterion of relevance is applied to the set of documents judged relevant, and that the criterion on which the decisions are made remains constant, at least during the period of interaction at the terminal. The assumption of consistency, as it relates to the employment of relevance as a performance criterion, had been the subject of an earlier criticism by Doyle -

> ''the most relevant subset'' is not only an individual matter for the searcher, dependent on the time and circumstances of his searching foray but also the feedback he gets is quite capable of changing his idea of what he wants as well as changing his way of expression. An 'information need' is thus revealed to be a dynamic entity, whose times of greatest dynamism and change may come in the very process of interacting with a retrieval system' (Doyle, 1963:199)

Analysis of the different criteria used in actual relevance judgements using a model interactive retrieval system INSTRUCT also highlights problems in the assumptions made about relevance judgement in work

on relevance feedback (Ellis et al., 1987: Nelis, 1985).

Nevertheless, for the purposes of the calculation of probability of relevance, relevance is treated as a boolean variable, the relevance decisions are assumed to be simple, independent of each other, and made according to the same criterion. The relevance judgements are also assumed to have an objective correlate in the document representations. This last point being interpreted as meaning that there must be a statistically significant similarity between the members of the relevant set, or a statistically significant difference between the set of documents judged relevant and those judged not-relevant.

These assumptions, although computationally necessary for the calculation of probability, result in the employment of a simplistic view of the nature of the actual judgements of relevance being carried out. If the relevance judgements made at the terminal are not static but dynamic; if the searcher makes relevance judgements not according to a constant criterion but with different criteria; if the order of presentation of documents or the nature of the other documents in the set alters the criterion employed, then the information used to calculate the probability of relevance will be unsound.

The Quantitative Approach to Relevance Judgement

In the use of relevance as a performance criterion, and in its employment to provide feedback, similar types of questions arise about the relationship between the assumptions made concerning relevance judgement in the laboratory, and the nature of relevance judgements made in real life. These questions have been examined by Oddy (1981), whose critique is particularly useful in that it covers points made by earlier critics of the Cranfield tests, but also highlights difficulties involved when relevance judgement is used both to provide feedback, and to serve as the criterion for evaluating performance.

Oddy (1981) considered that

> 'in order to establish a fruitful relationship between the laboratory tests and their hypothetical real life analogues, we must ask two questions:
>
> 1. Is relevance-based effectiveness safely separable from other performance characteristics for experimental purposes?

> 2. Is relevance in real life the same as relevance in laboratory tests?' (Oddy, 1981:168).

He argued that although the answer to the first question is not clear the second question has to be answered in the negative.

> 'relevance is dependent on three factors related to the user - perception, purpose and knowledge - which are causally closely related to each other, and subject to variation in the course of an interactive search. The picture of relevance decisions that we are obtaining is very different in nature from that included in test collections...What implications does this argument have for the results of laboratory tests? It is that, in principle, they are inconclusive' (Oddy, 1981:169-170).

Oddy was also critical of the procedure employed in experiments on relevance feedback, where relevance judgements made at the terminal were employed both to provide feedback and to evaluate the effectiveness of the search.

> 'Experiments on relevance feedback, which exhibit relatively outstanding effectiveness, make use of the same set of relevance judgements for two separate purposes. First, they are used to simulate the user's feedback, that is his reactions to documents retrieved, and thus they determine the query modification. Second they are used to evaluate the effectiveness of the technique. In real life, two distinct sets of relevance judgements would be used for a corresponding experiment. As he sits at the terminal, the enquirer would make instant judgements, according to his perception of the documents during the search session. His evaluation of

> the search would be made at a later time, on reflection. Theories are unrealistic in this respect, and experimental arrangements fulfil their unrealistic assumptions, and are thus inconclusive in relation to real life systems' (Oddy, 1981:170)

The very low number of items which were actually consulted following a search session, as opposed to being judged relevant during, computer based searches, has also been remarked on by Smithson (1987).

Oddy's (1981) critique provides a lucid analysis of the characteristic features of the relevance problem in information retrieval research, but it does not indicate why the problem has been so intractable. In order to explain this, it is necessary to return to the analysis of what was being done when relevance was adopted as a performance criterion, and why the approach to relevance judgement employed in work on relevance feedback is necessary to that work despite the recognised problems.

Cleverdon had correctly seen that if relevance was to be employed as a performance criterion the relevance judgements used would need to be set up in such a way that they seemed sufficiently objective to serve as the basis of measurement, and yet appeared to have a significant relationship with behavioural reality. A particular type of relevance judgement and associated procedure were then devised to control the type of judgement being made. This kind of control is not unique to the employment of relevance in information

retrieval research, in fact, it is a characteristic feature of scientific explanation that a concept with a large number of meanings (such as force, power, energy) is not employed directly in the theoretical structure, rather a measurable phenomenon substitutes for the everyday meaning and the term is now defined within the theory in terms of the operational measurement (Wittgenstein, 1967).

However, in the case of the employment of relevance as a performance criterion the situation is more complicated in that it also represents an attempt to formalise, for purposes of measurement, an aspect of human intellectual ability. This raises problems which are different in kind from those involved in the formalisation of an equivalent concept relating to some physical process. Cleverdon, however, seems to ignore this distinction. In the report of the second series of Cranfield tests, he quoted approvingly Wilkins (1964) remark that -

> 'There is a confusion of ends and means in this type of attack on measurement in principle. Perhaps if medicine threw away the thermometer, the encephalograph, the X-ray, and all the other technicalities, medicine would become much more human! How much more preferable the tender hand on the brow than a nasty piece of glass in the mouth - how inhuman! But is it sympathy and fellow-feeling that we want from the physician or a technical competence to identify the condition and give us the cure? The bedside manner still has a place in the cure, even although the hand on the brow has been replaced by the thermometer' (Wilkins, 1964:9 in Cleverdon and Keen, 1966:192)

But it is not objection to measurement which underlies criticism of the employment of relevance as a performance criterion, rather, the lack of recognition that there is a difference in kind between the employment of a device to measure a physical process, and the employment of human judgements of relevance in tests of the effectiveness of information retrieval systems. The use of peoples' judgements of relevance creates a situation where the nature of the problems of measurement are different in kind from those involved in the use of a physical measuring device. This difference in kind is illustrated by a remark of Putnam that

> 'it is a feature of 'scientific' knowledge (at least if we take physics as the paradigm) that we use measuring instruments that we understand. Our theory applies to our measuring instruments and to their interactions with what they are used to measure, not just to the objects we measure. It is a feature of practical knowledge that we often have to use ourselves (or other people) as the measuring instruments - and we do not have an explicit theory of these interactions' (Putnam, 1978:72).

The employment of human judgements of relevance in information retrieval testing relates to the second type of 'measurement', and the difficulties which arise in the use of relevance in information retrieval research derive from the quantitative approach to the notion necessitated by its employment as the criterion of measurement. In the case of the work on relevance

feedback a quantitative approach to relevance is necessary both because the notion is used as a criterion of measurement, and because it serves as the basis for deriving the information on which the calculation of probability is made.

For the purposes of this study, the influence of the assumptions on the direction and nature of the research carried out is more significant than their effects on the actual results of the test themselves. The quantitative approach to relevance, though necessary for the methodological and computational requirements of this tradition of research, historically led to comparatively little interest being shown in the use of qualitative approaches to relevance judgement in information retrieval research, despite the availability of models which could have served as the basis for such development (Kemp, 1974: Wilson, 1973).

In 1963 Doyle, considering the role of relevance in information retrieval tests, had concluded that -

> 'Relevance will serve its purpose, but will decline as the realisation slowly comes that an individuals information need is so complex that it cannot accurately be stated in a simple request...The gradually increasing awareness of a human's incapability of stating his true need in a simple form will tend to pull the rug out from under many information retrieval system evaluation studies which will have been done in the meanwhile' (Doyle, 1963:200).

Doyle considered that the solution to this problem was not to design systems around the concept of relevance but to base them on the concept of exploratory capability. Doyle argued that

> 'the searcher needs an efficient exploratory system rather than a request implementing system' (Doyle, 1963:199)

and that

> 'Exploratory capability, as it turns out, is provided by traditional libraries, but not by some of our modern machine literature searching schemes' (Doyle, 1963:199).

This type of approach to information retrieval system design, requires rejection of the simplifying assumptions of the information retrieval model, and the associated quantitative approach to relevance. Instead, the various ways in which the interaction of the searcher with the system might take place needs to be examined or modelled, and appropriate design configurations evolved for the intended context of use. The features of two different generic approaches to the characteristics of users studied, and to the form of employment of those characteristics in information retrieval system design will be outlined in the following sections.

Chapter Two

Cognitive and Behavioural Approaches to Information Retrieval System Design

The Cognitive Approach

In recent work on information retrieval system design the major alternative to research employing the information retrieval model has been concerned with the development of cognitive models for information retrieval. Reviewing this work Daniels (1986a)

identified three types of cognitive models, mental models, conceptual models, and user models, and distinguished between two types of user models: empirical quantitative models and analytical cognitive models (Carbonell, 1983: Daniels, 1986a). Most studies adopting a cognitive approach to information retrieval system design have been concerned with developing analytical cognitive user models (Daniels, 1986a). For this reason, in this study, it is work related to the development of these types of models which will be referred to as representing the cognitive approach.

Two key features characterise this approach to information retrieval system design (Hollnagel and Woods, 1983).

1. The construction of a model of the user in the system,
2. and the derivation of this model from cognitive characteristics of the user.

Daniels (1986a) gives examples of a large number of different types of cognitive user models operating in various application areas including help systems, expert systems, computer assisted instruction, and intelligent front ends.

However, although the development of cognitive models for information retrieval systems design has aroused considerable interest (Belkin, 1984: De Mey, 1980: Hollnagel, 1984: Ingwersen, 1984: Pejtersen, 1984: Wormell, 1984), comparatively little work has

been carried out employing such models in information retrieval research - and the work which has been undertaken has had a number of different objectives, and often only a limited empirical basis. In some studies it is not made clear how extensive the empirical component is (Belkin, 1984: Belkin, et al., 1983), and the empirical content of many of the others has been relatively slight. In one series (Brooks, 1986: Daniels, 1986b: Daniels, et al., 1985) the core data set consisted of only six transcripts, and Ingwersen's (1982) work was based on analysis of the interaction between thirteen librarians and only five users.

The major studies undertaken have been concerned with attempts to represent the problem situation (Wersig, 1971: 1979: Wersig and Windel, 1985), with attempts to represent search strategies (Ingwersen, 1982: Pejtersen and Austin, 1984), with automating the role of the expert search intermediary (Belkin, 1984: Belkin et al., 1983: Brooks et al., 1985: Brooks, 1986: Daniels, 1986b: Daniels et al., 1985), on the development of an expert assistant for document retrieval (Croft, 1985a: 1985b: 1986), on user modelling via stereotypes (Rich, 1979: 1983), on information retrieval through man machine dialogue (Oddy 1977a: 1977b), and on the development of anomalous states of knowledge based systems (Belkin, 1978: 1980: Belkin et al., 1979: 1982a: 1982b).

Not all of these studies have an appropriate focus or are sufficiently empirically based to be of use here. The suggestions for basing an information retrieval system around the notion of problem situation (Wersig, 1971: 1979: Wersig and Windel, 1985) seem to have advanced little beyond the stage of abstract speculation and have been the subject of considerable critical comment (Ellis, et al. 1985). Some of the studies have been solely concerned with fiction retrieval (Pejtersen and Austin, 1983: 1984). The attempts to represent search strategies (Ingwersen, 1982), and the work undertaken on developing a system for automating the role of the search intermediary (Belkin et al., 1983: Brooks, 1986: Brooks et al., 1985: Daniels et al, 1985: Daniels, 1986b), although interesting, relate only to an aspect of the information retrieval system interaction, and represent a rather limited view of the contribution of user modelling to information system design, being restricted to analysis and modelling of the interaction between searcher and search intermediary in the reference interview.

The most ambitious in conception, though incompletely implemented and tested, of the attempts to develop an automated expert search assistant has been that of Croft's expert intermediary system (Croft, 1985a: 1985b: 1986). The expert intermediary system is intended to provide assistance to the user throughout

the retrieval interaction by means of a number of 'system experts'. These Croft (1986) has characterised as the User Model Builder, the Request Model Builder, the Indexing Expert, the Thesaurus Expert, the Search Controller, the Browsing Expert, and the Explainer. However, the expert intermediary system itself is not yet at the prototype stage, as Croft points out,

> 'The basic architecture of the I3R system has been implemented although some of the experts are in skeleton form. A scheme for representing and manipulating plans in the scheduler is currently being tested. The user interface for specification of domain knowledge has not yet been implemented' (Croft, 1986: 202).

Daniels (1986a) also notes that

> 'Croft's Expert Assistant has yet to be fully implemented and it is therefore difficult to comment on the adequacy of the user models in the system. However, they appear to contain very little detailed knowledge and seem to be confined to rather limited user descriptions, e.g. expert/novice' (Daniels, 1986a:297)

However, Croft (1985a) does point out close affinities between the expert intermediary system being developed and work on information retrieval via stereotypes (Rich, 1979: 1983), and on information retrieval through human computer dialogue (Oddy, 1977a: 1977b). Therefore, examination of these studies may help to illuminate the nature of the system being proposed, and to highlight difficulties which might be encountered in

its implementation.

In the work on user modelling via stereotypes (Rich, 1979: 1983), a system called GRUNDY was developed and tested which recommended novels to individual users on the basis of a number of stereotypes built into the system. Examples of such stereotypes were 'Any Person', 'Man', 'Woman', 'Educated Person', 'Sports Person', 'Intellectual', 'Feminist'. The stereotypes were activated by responses to characteristics of the users requested by, or presented to, GRUNDY in the course of the retrieval interaction. The characteristics were positively and negatively weighted with reference both to the stereotypes and to a set of novels. Based on the weights assigned to the characteristics, which were computed in the course of the retrieval interaction, GRUNDY proceeded to recommend certain novels to the user. Rejected suggestions were employed to re-calculate the weights in an iterative process until a suggestion was accepted or GRUNDY gave up.

The performance of the system in the task was good. On the basis of users' judgements of good and bad suggestions GRUNDY consistently outperformed random selection of novels (Rich, 1979: 1983). But, the study was carried out in the context of fiction retrieval, and it is questionable whether the same approach, using a limited set of stereotypes, would represent a satisfactory basis for the design of an information

retrieval system. The limitations of the stereotype approach in an information retrieval context stem from the fact that the stereotypes apply to the users of the system not to the problems they are dealing with. So, for example, if stereotyping were attempted on the basis of disciplinary boundary, or affiliation, this would not take account of the topics on which information was required, the complexity of perceptions of the nature of the topics studied, and the variety of different approaches to particular topics, which often transcend disciplinary boundaries. These make stereotyping even by this criterion potentially stultifying.

A few examples of how social scientists describe their research interests may help to illustrate the point that stereotyping social scientists by disciplinary affiliation would not adequately reflect the differences of approach to problems adopted by individual social scientists. The comments are from researchers working in the same Unit, with similar disciplinary backgrounds, and working on closely related research projects.

> 'I described my interests in very concrete terms - the psychological effects of unemployment - but there are different ways into that. I could have described them in terms of psychiatric epidemiology, in terms of theoretical models of human development, and so on. And depending on the way you describe it you bring in more and different kinds of people' [MRC/SSRC SAPU 1A 350]

'You join different bits of theories it's not a unified theoretical area. That is why it's impossible to say who is doing any work at one moment in time - you do get economists, labour market economists particularly, sociologists, social policy/social administration people, psychologists, doctors - a whole range of people doing different bits of studies' [MRC/SAPU 1A 367].

'Other relevant type of work could be on similar processes which have been studied in different areas - for example how people react to the loss of a husband or wife - or serious injuries - reactions to these events could be seen to be relevant or similar in their effects to that of losing a job. So that kind of work is probably worth looking into as well. I've read a certain amount about reactions to serious injury and grieving. What this is is looking at situations which are comparable and seeing whether there is any logical reason why people might react or respond in a certain way to a particularly stressful event - it may be that unemployment is not that different from any other stressful life event' [MRC/SSRC SAPU 6B 083-100]

To try and avoid such problems by using stereotypes of problems rather than users seems fundamentally unsound. The reason being that it is not possible to enumerate a limited set of stereotype problems with the related information concerning those problems. Problems are not pre-defined, and decisions about what information relates to which problem is a part of problem definition. This is a dynamic not a static activity, as one of the researchers noted -

'Every now and again I go over the old issues

> of the journals and I think I wonder why I didn't read that before it really is quite important. Realise that somebody's thinking was way in advance of mine - or alternatively parallel and irrelevant at the time - they couldn't see a link but I can see it now. The point is I should have read it and absorbed it at the time, or rather, it would have been useful to me now if I had done' [MRC/SSRC SAPU 4A 2341]

Nor are problems self-contained, and again, decisions concerning the boundary of relevance of information to a particular problem are a part of the person's definition of that problem. This is not fixed and may be affected by the person's interests or theoretical standpoint, the current influences on the field in which the problem is located, or the level at which the problem is being treated. Examples from each of the groups described later may help to illustrate the complexity of these issues in relation to actual problems experienced by social scientists, and highlight the difficulties involved in attempting to stereotype problems.

> 'There's a big connection between linguistics and computer science - there's an Association of Computational Linguistics. But they spend most of their time trying to make computers understand natural language for example. Whereas, what we are doing is applying the same kind of notions of models of knowledge that linguists have to the problems of human computer interaction, and I think that this is probably quite different - I would say it was unique - an unusual theoretical approach to cognitive ergonomics' [MRC/SSRC SAPU 9B 555]

'The nature of theoretical explanations in psychology is showing some signs of very drastic change - principally I think as the result of the impact of artificial intelligence techniques and theoretical approaches. These are the subject of some quite topical debate amongst psychologists. And that again opens up the necessity to try and get abreast of an initially unfamiliar literature or that selection of it that can be grasped by a non-AI specialist and brought to bear on his or her particular problems' [Psychology 1A 171 181]

'There are groups of people who create a paradigm which they explore using different methods, for example, that they are keen on. Another paradigm will use a qualitative approach and wouldn't be so interested in quantifying particular factors. They would want to pursue meaning, and they would want to pursue understanding, and so on, and that would be in terms of social science of education, quite a dominant paradigm, against which school and teacher effectiveness might be placed' [Education 3C 043]

'The amount of hard data which there is on the difference that locally based social work makes is minimal. The amount of work being done in the area is not vast. Therefore, that initial read round could be covered fairly easily. But once you go the next level of what is relevant you then encompass an enormous range of stuff, because you're then concerned with general theories about social work. You're also concerned with the whole - the sociology of community - on which there is an absolutely colossal literature...We are concerned with the health and social services. We will therefore read all the literature about district nursing - about health visiting - about housing wardens - about educational welfare officers - general practitioners - social workers - all of those are subjects in themselves. So as well as the whole question of sociology of community there is also all the detailed literature

> about all these worker groups. Also professionalisation - the sociology of organisations which is an implicit part of it - and the sociology of the professions' [Sociological Studies 10C 260-304]

Where stereotypes may have a contribution to make to retrieval system design might be in relation to the 'expert/novice' user distinction already mentioned, and perhaps in relation to 'theoretical/experimental', 'quantitative/qualitative', 'survey research/case study', type distinctions. Apart from this, restricted, form of stereotyping, and, although interesting, particularly in relation to fiction retrieval, the use of stereotypes for information retrieval purposes has severe limitations.

The work carried out on information retrieval through man-machine dialogue (Oddy 1977a: 1977b), and on the development of an information retrieval system based on the concept of anomalous states of knowledge (Belkin, 1978: 1980: Belkin, et al, 1979: 1982a: 1982b). are of particular interest for this study, since both approaches are directed to information retrieval system applications, both carry implications for database construction, and both have been the subject of some empirical research.

In the work carried out on information retrieval through man-machine interaction Oddy (1977a: 1977b) developed an information retrieval program referred to as THOMAS. THOMAS is described as an information

retrieval program rather than an information retrieval system because, as Oddy stressed, in itself THOMAS could not provide

> 'a comprehensive information facility for the researcher' (Oddy, 1977a:12).

THOMAS operated through a series of interactions between man and computer. It did not employ a request statement nor was relevance information employed in quite the manner characteristic of relevance feedback experiments though relevance information was used in the program. The principle behind THOMAS was that

> 'the machine forms an image (rather as a man does) of the view of the human enquirer, without requiring him to ask a precise question, and responds with references according to its image' (Oddy, 1977a:1)

The effectiveness of the THOMAS program is crucially, and necessarily dependent on the quality of the image it can build of the searcher's information requirements and, in turn, this is dependent on the quality of the interaction between the man and the program. From the interaction with the searcher the program

> 'forms an image of the searcher's interest derived from its world model, chooses references for display according to the state of the image, and modifies this image continuously in the light of his reactions to displays. The program's world model consists solely of knowledge about the organisation of

> literature. Specifically, it is a network of associations between documents, authors, and subject terms: any pair, of like or differing type may be linked. In principle any other type of entity of interest to a literature searcher may be represented in the network' (Oddy, 1977a:5).

So the database is structured on objective properties of the material in the database - with associations between these properties being of particular significance. Interestingly,

> 'no information is held on the network of associations: it is sufficient for this program to know simply that two entities are associated' (Oddy, 1977a:5).

The program then proceeds by an interactive procedure to attempt to identify documents most highly associated with the criteria that the searcher appears to employ in decisions taken during interactive sequences with displays of document representations. At no point is the searcher expected to build up an image of what the machine is doing, of what image the machine is operating with, of the nature and manner in which either the program or the database is structured, or of what the program is doing with the database in the course of the interactions. Though Oddy notes that in the course of one dialogue -

> 'The user had noted that the term 'time factors' was attached to several of the references, and had a wide variety of

> meanings, so he now stated that he was no longer interested in it (here is implicit evidence that the searcher is also building a model, or image: to represent the program's view of the world)' (Oddy, 1977a:10)

- it is clear that this is almost incidental. The THOMAS program itself had nothing in the way of features that might explain what the program was doing and the sort of information that it had available, in terms of its own representations, to work on. This places a high degree of confidence, for effective retrieval, in the quality of a dialogue which Pinsky (1983) and Nelis (1985) show is fraught with difficulties.

The THOMAS program could be enhanced to provide more rapport (Cheydleur, 1961) between the system and the user. But such enhancements would leave the division of labour between man and computer unchanged. The program would still construct the image - which might or might not be an appropriate one for the user. If the associations the searcher makes are accurately reflected in the associations recorded in the database then the program may work effectively - if they are not it is difficult to see how it can work at all.

Belkin's (1978) idea for an 'anomalous state of knowledge' based retrieval system connects up with Oddy's ideas at this point. For Belkin's proposal also relies on a database consisting of a collection of items and their associations, and the interaction of

the system with the searcher as it attempts to model the searcher's anomalous state of knowledge has clear affinities to the kind of interaction which underlies Oddy's program. In Belkin's scheme

> 'A recipient instigates the communication system by recognizing an anomaly in her/his state of knowledge, this recognition being akin to the partition of generator's state of knowledge which identifies the conceptual structure to be communicated. The recipient then converts this anomalous state of knowledge (ASK) into some communicable structure (e.g. a request), which is used to retrieve from the corpus of texts some other text or texts which might be appropriate for resolving the anomaly. The recipient interprets the text to discover the conceptual structure underlying it, this structure interacts with the recipient's ASK, and the recipient then makes a decision as to whether the anomaly has been sufficiently resolved.' (Belkin, 1978:81).

A design study on the feasibility of employing ASKs as the basis of information retrieval was undertaken, (Belkin, et al., 1982a: 1982b). but this was limited to the derivation of the ASK representations. There was no attempt to search any database with the ASK representations, so it is not possible to assess difficulties involved in matching the ASK representations with document descriptions or to evaluate the technique in terms of any criterion of retrieval effectiveness. The results of the design study are interesting in terms of the procedures for creating the term association networks from problem descriptions and from abstracts. But, in the absence of

any clear statement as to how the matching process would operate, or any attempt to employ the technique for retrieval, it is difficult to evaluate how far, in practice, it differs from other term association or term clustering algorithms.

Nevertheless, the study is worth noting since it explicitly recognises the similarities between the ASK approach and the THOMAS program and attempts to use these as the basis for design recommendations. The affinity between the two and their integration in the cognitive approach is stated clearly in the background paper to the design study.

> 'Thus, the cognitive view of the IR situation leads us to some design principles for IR systems which we think incorporate the THOMAS and ASK approaches. These design principles require representation of the user's anomalies, evaluation in terms of the problem the user faces and iteration and interaction in retrieval' (Belkin, et al., 1982a:65-66)

The system would be interactive in something like the same way that Oddy's THOMAS program was, in that the system would attempt to construct a model of the searcher's requirements from building up an image of the searcher's state of knowledge and comparing that to the state of knowledge held in the system.

The ASK hypothesis itself has been the subject of criticism (Moser, 1978: Pratt, 1978). The fundamental criticism which can be levelled at the ASK concept is that it embodies the term 'knowledge' but makes no

reference to any propositional or logical context, nor is there any reference to semantics, norms, rules of evidence, what constitutes a valid explanation, a properly carried out experiment, or whatever. In the ASK concept terms substitute for concepts, and term networks take the place of knowledge structures. The design study makes explicit reference to the substitution of word, or term, associations for concepts, justifying this in terms of ease of data analysis.

> 'The formalism used for representation, a network of concepts represented by words, depends upon our concept of a state of knowledge as a multi-dimensional structure...the basic idea seems to us unexceptionable, except perhaps for the substitution of words for concepts. A network is certainly an effective way of representing this type of structure, and has the advantage, in our case, of being relatively easily derived from the sort of association database we use' (Belkin, et al., 1982a 68)

However, there is no justification, other than that of computational convenience, for equating term associations with knowledge structures. Whatever the structure is which underlies Belkin's argument and system, it is not a knowledge structure.

In the absence of an argument to equate networks of term associations with knowledge structures there seems little to distinguish the structure of a database based on ASK association networks from one which was based on simple term associations such as that employed

by Salton (1971: 1973). Even if the ASK hypothesis had taken into account the kind of considerations that might have made it appropriate to refer to the concept as being concerned with knowledge structures, it is difficult to conceive of how this could be useful for subject areas where the question of what is knowledge and what is not is fluid or subject to dispute. In such cases the very idea of what represents an anomalous state of knowledge becomes problematic.

However, whatever the epistemological status of Belkin's concept, and despite the questions concerning how the matching process would be implemented, there are evident similarities with the framework and the procedure which underlies Oddy's THOMAS program. Both require that the system or program build a form of cognitive model of the requirements or perceptions of the searcher wishing to retrieve information from the database as a prerequisite for retrieval. In both cases, the database of is structured by term, or other forms of association between items in the database, and the retrieval system or program is expected to build some kind of model or image of the searcher's requirements or perceptions which is then matched against the database structure. There are further similarities in the distribution of tasks between man and computer in the retrieval interaction. Therefore, for present purposes, despite differences of style and emphasis between the two, they can both be considered

as being similar in approach.

The Cognitive Approach and Characteristics of Social Science Information

In the cognitive approach to information retrieval system design the searcher does not deal directly with the database as such, he interacts with the retrieval program which then interacts with the database, via the creation of a model of the searcher. The model which the system constructs of the searcher's requirements, or state of knowledge, is then compared to the database, so that the building of the model is a necessary pre-requisite for retrieval. In principle, any type of associations in the database could be utilised to build the model, in practice terminological associations play the major role. These are employed both in structuring the database, and in modelling the

searcher's requirements or state of knowledge. The question to what extent the model which a program or system of this sort can properly be said to be building up a genuinely 'cognitive' model will not be considered here. Instead, the practical difficulties involved in creating a model of this nature in the context of designing a retrieval system for academic social scientists will be addressed.

If terminological associations play a large part in structuring the database or in the effort to build up a model of the searcher's requirements, or state of knowledge, severe problems may be experienced in the context of social science information. The particular difficulties relating to terminology in the social sciences have been the subject of frequent comment (Adam, 1982: Ellis, 1982: Thouless, 1939: Wersig, 1970: Wilson, 1980: Wright, 1973). The problems identified have included, lack of definition of the field, use of everyday words with technical meanings, variation in the use of terms, changes in meaning of technical terms coming into common usage, and the use of the same term with different meanings from different ideological or philosophical perspectives. It would not be correct to assume that problems with terminology were unique to the social sciences but they do appear more prevalent and difficult to deal with than similar terminological problems in the sciences (Foskett, 1974).

Difficulties with terminology in the social

sciences have led to concern with, and occasional proposals for, the standardisation of terminology. In the case of psychology, this can be traced back to the turn of the century (Claparede, 1910: Ruckmich, 1913: Tonnies, 1899: 1900: Warren, 1911: 1912: 1913: 1916: 1918: 1920: Warren, et al., 1918: 1922: 1925). A more recent attempt at standardisation of social science terminology was the Interconcept project undertaken by UNESCO (Dahlberg, 1978a: 1978b: 1980: 1981: Riggs, 1979a: 1979b: 1979c: 1979d). However, efforts at standardisation of social science terminology seem to rest on a basic misconception of the nature of the social sciences, as Winn (1971) observed -

> 'It is all too easy to make assumptions about social sciences by analogy with sciences. It is a truism to note that natural scientists have agreed ways of categorising and describing the phenomena in which they are interested. The indeterminate boundaries of social science fields have often been the subject of comment (unfavourable comparison). This lack of 'agreement' and indeterminacy are however a natural consequence of the fact that a social science discipline does not deal with a 'special class of empirical data; instead it deals with data as interpreted within a special type of conceptual framework' (Smelser, 1969)' (Winn, 1971:81)

That terms change their meanings and that the same terms can refer to different things from different perspectives is part of the nature of enquiry in the social sciences, as is the two way interaction between the terminology of the social sciences and everyday

language. This may affect some social science disciplines more than others -

> 'Such problems are particularly marked in the field of sociology, where meanings of terms change not only as sociology develops but as society itself changes' (Watson, et al., 1973:275)

But the point applies in some measure to all social science subjects because of the nature of the subject matter the social scientist studies.

> 'Social science concepts cannot be seen as corresponding to a single referent because they cannot be dissociated from their everyday meanings and the everyday life which forms the subject matter of the social scientist' (Watson, et al., 1973:278)

Difficulties with terminology were mentioned by several of the social scientists interviewed in relation to problems experienced in formulating a search request for computer based searches, these will be detailed later, and are similar to reports of other studies of computer-based searching by social scientists (Hounsell, et al, 1977: Ellis, 1982: 1984a). Of particular interest are the comments of one educationalist, who compared his experience in education with his previous experience in physics. He considered that there were particular problems with terminology in the social sciences which had no equivalent in the sciences, and that the difficulties

of this nature he had experienced since coming to work in education had no correlate in his previous experience in physics.

> 'The thing about physics was the clarity of definition of the area you were working in. If you were looking at the abstracts you could go immediately to the small section which were specifically in the area you were interested in. Whereas, it's the nature of education, and I suspect social science research generally, that descriptors are hard to find. Terminology is not so tight. If you're working in geophysical research, on some small branch of it like ionospheric physics, everybody knows, indeed in other languages, that they use the same words' [Education 5D 372-385]

In contrast, in education,

> 'its much more multivariate, if you like, and much less agreed across people about the meanings of words, or even whether the word is appropriate to describe things...Using the same words but actually meaning quite different things by them' [Education 5D 405-439]

He considered that what was behind this difference between the sciences and the social sciences was a difference in the nature of the concepts employed by social scientists compared to those employed by scientists.

> 'I think that the concepts you are dealing with are hypothetical in a quite different way to scientific concepts. They are hypothetical - I mean you don't see a molecule, you don't see a bond between

> molecules, you don't see an ionosphere either - but the evidence for them is reasonably tight. If you talk about something like a misconception in science there is a whole body of interpretation between the actual phenomena, what you use to describe it, and what you use to try to explain it. The major difference it seems to me is the degree of closure that you can bring to bear on what could be an appropriate theory. When Newton said that he did not create hypotheses he was talking about the sort of hypotheses you get in education rather than the sort of hypotheses you get in science. The whole notion of a misconception, for example, has an enormous cluster of unobservable things which go between someone having attempted to teach a child about an area of experience and the child demonstrating, through behaviour, something which is interpreted as being a different conception of the world from what was taught. Now that bristles with hypotheticals in a way that a reflection of a radio wave from a part of the upper atmosphere does not. There is a great imprecision in the whole notion of what cognition may be, and between the relationship between experience and cognition, and between that and observed behaviour is unknown. So that the whole linkage between contrived experiences in schools and behaviour of children is so tenuous and unknown that a wide variety of different hypotheses can be put forward to attempt to explain the the same thing, and the evidence may not be tight enough to distinguish between them' [Education 5D 456-506]

If the terms which social scientists use bear a different relationship to the knowledge structure of the social sciences compared to that which they do in some sciences, then terminology alone may provide a better basis for designing a system for science subjects than it does for designing one for social science subjects. In the context of social science information, identifying that part of the database

which will be of most interest, or even making any sense of the searcher's requirements or knowledge structure, on the basis of term associations, may be difficult or even impossible.

Social scientists often make decisions on whether to follow up material not merely on the basis of its substantive topic but also on the basis of the theoretical perspective or the tradition of work that has illuminated it. The decisions may stem from perceived differences in theoretical assumptions, methodology, ideological persuasion, or assumptions as to the sorts of problems that are thought worth examining. Issues of this nature have been raised in considerations of the question of the appropriate form of approach to index language design for social science subjects, in particular for sociology (Hutchins, 1978: Swift et al., 1974: 1977a: 1977b: 1978a: 1978b: 1979).

There has been strong criticism of the failure of information services to recognise and reflect the differences of perspective encountered in a subject such as sociology (Watson et al., 1973). These issues are related to the kind of terminological difficulties already considered -

> 'Problems of terminology are apparent not only when an individual attempts to cross discipline lines but also when he wishes to locate his problem via the terms of another theoretical viewpoint within his own discipline' (Watson et al., 1973:275)

This can create particular difficulties when employing an information retrieval system.

> 'For example, when two documents are put into the same category because the same term occurs in both then users of the system have to assume, along with the creators of the system, that the documents are dealing with a common entity to which the word refers. But, as most sociologists would agree, this is to oversimplify the nature of many, if not most, terms in sociology. The meanings of terms are mediated by the theoretical perspective within which they are embedded. Understanding of the perspectives in a discipline is acquired through communication and the development of shared understandings, which are, in part at least, implicit' (Watson et al., 1973:278)

Nor does reliance on terminology deal adequately with the problem that different groups may systematically operate with a different understanding of what constitutes the scope of the subject of study - this was noted by one of those interviewed in relation to the subject of thematic mapping.

> 'I'm coming to the view that the American definition, the North-American definition, is different perhaps from the British definition, and that we would define the term more widely than the Americans who have a rather more abstract definition...UK workers define 'thematic' more widely and therefore expect a wider coverage, whereas the operative definition...used there is a fairly narrow one' [Geography 6C 076]

To cater for these differences by relying on the system constructing a cognitive image or model of the

searcher's perceptions, or of his state of knowledge, is an ambitious and complex task, particularly if the searcher's image is understood as being dynamic rather than static, as another of those interviewed pointed out -

> 'You've obviously got some some kind of model in your head, and some kind of relationships, and all you can do is to identify with those relationships - and if you find the model shifting all the time then what you're looking for is shifting all the time' [Economics 3D 086]

Finally, there are intrinsic difficulties involved in the nature of the interaction which has to take place in order for the program or system to build up an image of the searcher's requirements (Green, 1983). If the dialogue between the user and the system is defective then any information derived from that interaction which is subsequently employed by the program or system will be deficient or misleading. Pinsky (1983) and Nelis (1985) describe examples of breakdowns in communication between computer systems and human users: these would make salutary reading for those who place great faith in the ability of computerised retrieval systems to build up complex models of human cognition on the basis of human computer interaction.

The Behavioural Approach

In the context of designing a system for social scientists, the cognitive approach, particularly in the form taken by Oddy and Belkin, seems to face serious obstacles, both in the derivation of the cognitive user model, and in its employment in an operational information retrieval system. These have their source in the attempt to construct a cognitive model of the user on the basis of his interaction with the system, and in the reliance on terminology to derive this model. This does not invalidate the idea of employing some form of user modelling in information retrieval system design, but does raise doubts concerning this approach to the construction and use of such models.

The major postulates of this study are that user modelling for information retrieval system design does not require the construction, by the system of a cognitive model which mirrors the complexity of the searcher's cognitive world, and that the construction of a model of the user's requirements, or state of knowledge, as a pre-requisite for information retrieval is otiose. A different approach to user modelling for information retrieval system design is possible, based

on study of what people do in the course of their information seeking activities, and of the perceptions underlying these activities. The activities and perceptions - which together can be understood as constituting the individual's information seeking pattern - can then be broken down into their characteristics, and the characteristics be used as the basis for system design.

It should be borne in mind that these characteristics are not intended to represent a set of prescriptive search heuristics of the kind outlined by Bates (1979a: 1979b: 1981) or Markey and Atherton (1978). The derivation of the behavioural model and its application to questions concerning information retrieval system design, represents an attempt to provide a comprehensive analytical model for describing actual information seeking behaviour, which relates both to the extensive body of research on user studies, and to research carried out on human-computer interaction in an information retrieval context.

To illustrate the nature of the analysis by which the information seeking characteristics can be 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 who 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 SAPU TG 3024]

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.

1. Starting: identification of a key paper to commence the search.

2. Chaining: following up references to this paper, and following up book advertisements from the journal consulted.

3. Browsing: to identify relevant journal sources.

4. 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 4 or 5 journals that may well have something of

note. The main source is 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 3 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 a 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 041-071]

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

1. Monitoring: maintaining awareness of developments in an area through regularly following particular sources.

2. 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...Its 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. Its only when you start looking for something that you find them' [Education AWE 041-091 102-118 125-127]

This aspect of the pattern shows how the different characteristics of the pattern interlink and, to a large extent, complement each other, as the 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, also 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 5 years Economic History Reviews, which are standard, that is the economic historians' journal - Economic History Review. I looked through that for the last 5 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. Its the obvious ones to start with' [Economic and Social History DEB 451-473]

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).

From the analysis of the information seeking activities of all the social scientists interviewed six generic characteristics were identified.

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 areas of potential interest.

4. Differentiating: using differences between sources to filter the nature and the quality of the material examined.

5. Monitoring: maintaining awareness of developments through monitoring particular sources.

6. Extracting: systematically working through particular sources for material.

These six characteristics seemed sufficient to provide a detailed and comprehensive behavioural model for the design of an information retrieval system for academic social scientists. In the following sections the features of this model are described in more detail, and recommendations for the development of an information retrieval system based on the model are outlined.

Chapter Three

The Behavioural Model: Overview

The behavioural model has been derived from analysis of the information seeking patterns of all the social scientists interviewed. The overview of the model outlined here is cross referenced to the individual responses of the social scientists, which are described in the following chapters. Essentially the features of the model have been derived from analysis of the interviews, so that the characteristics are grounded in the responses of those interviewed

(Glaser and Strauss, 1967), however, where appropriate, reference is made to the findings of other studies of social science information use. Following the description of the information seeking activities of the different groups in terms of the behavioural model, the model is then used as the basis for an analysis of the features which might be required for an exploratory information retrieval system for use by academic social scientists.

The model is applied to four different broad groupings of the social scientists interviewed. The groups consisting of those interviewed in -

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

This analysis is intended to highlight variations in detail of the model in the contexts of the information seeking activities of the different groups, while serving to underline the general applicability of the model to those groups.

Starting

Starting refers to characteristics of the information seeking patterns of social scientists who are commencing work on a new topic or in a new area. The person concerned may be experienced or inexperienced in research, and may have some or no familiarity with the topic or area. Three generic ways of starting are through the use of

1. starter references;
2. reviews or synoptic articles;
3. and secondary services.

These three methods can interlink with each other. A starter reference may take the form of a review or synoptic article, and starter references or reviews may be identified from a preliminary search of a secondary service.

Starter References

A starter reference is something which can serve as a starting point into an area and from which other ways of gathering material can be established. The use of starter references whether previously collected or newly recommended allows the social scientist to get some purchase on a new subject, but at the same time to do that in a a way which allows other characteristics of information gathering to be quickly established. Often what is being looked for at the start is something embryonic which can be developed and added to as time goes on.

Different means are employed by social scientists to obtain such starter references. Many of the social scientists interviewed said that their first step 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. The employment of informal contacts is frequently mentioned in studies and reviews of social science information use (Adam, 1975: American Psychological Association, 1963: 1965: 1969: Brittain, 1970: Cronin, 1982: Hogeweg-De-Hart, 1983: Line, 1971) and occurs as an aspect of several of the

characteristics of the social scientists information seeking patterns.

The more experienced researcher having to start off on a topic about which he knows comparatively little has the advantage over the novice in this respect, in being more likely to have such contacts, but even in cases where people did not know anyone in the new area they could still pursue means of searching for information which involved establishing such contacts. An advantage of such an approach is that personal contacts can be employed both to provide references, and evaluations of quality or importance, thus allowing the person to concentrate on what are perceived to be key references, or key ideas, and to work from those. [Psychology 6A 2371: Continuing Education 10D 360-367: Economic and Social History 4C 566: Sociology 10C 515; 13C 2048-2070].

Many of the social scientists interviewed had some awareness of a broader area of topics than those which they were directly working on at a particular time. In such cases they often already had references to key people or material, and their natural tendency was to begin with what they were aware of and to use this as their starting point, relying on what they already knew and commencing the search from there. Sometimes individuals made explicit provision for this and maintained files of such references. Less explicitly, they made a note of references on topics tangential to

their principal interests, these then become potential starter references. [MRC/SSRC 1A 650; 8A 250-260: Psychology 4A 457: Education 4D 366-384].

Students and novice researchers are likely to be particularly dependent on supervisors or colleagues providing material to start them off on their research. Some of the social scientists interviewed, who had responsibility for supervising research students or research assistants, made explicit reference to this aspect of their supervisory role. Others who were asked what they thought they would do if they had to induct a research student or research assistant into their area, gave a similar response to those who actually had experience of supervision. [MRC/SSRC 7A 694; 12A 577: Education 6D 573; 7C 481: Economics 1C 2116: Economic and Social History 5C 184: Sociology 10C 207-234]. Developing an awareness of the different approaches to a topic, or induction into a particular perspective, was also mentioned as an important feature of starting for novice researchers. Though the exact manner by which this was done varied between the 'liberal' and the 'catholic'! [Psychology 1A 565; 3A 321-360].

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. There is some evidence to suggest that the ratio of review articles to primary articles is far smaller in the social sciences compared to the sciences (Bath University, 1975), but this may be compensated for by reviews contained in monographs (Line, 1976), as monographs are relatively more likely to be perceived as representing a main information source by social scientists than by scientists (Skelton, 1971).

Skelton (1971) also concluded that reviews were not perceived as particularly useful for locating information by either scientists or social scientists. Nevertheless, reviews or review type material may be found particularly useful when starting since these not only provide a source of reference to primary material but also a context or framework for understanding that material. Reviews were particularly singled out by the psychologists interviewed as representing a convenient means of getting started in a new area, particularly if

the area was characterised by a well established conceptual structure and a large body of empirical research. [MRC/SSRC SAPU 8A 275 + 364-371; 1A 2145; 5D 490-510; 7A 2011-2027; Psychology 4A 169].

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. [12A 2400-2414: Economic and Social History 4C 462-473]. Synoptic tutorial articles may also be understood as representing a form of review, such articles can be especially valuable in fast moving areas which are frequently characterised by rapidly changing conceptual structures. [MRC/SSRC 4A 3122-3131].

Also of value when starting are collections of papers, which, although not reviews in themselves, can together provide an overview of an area, and can represent an excellent way of starting. [11A 2099-2169]. They have the additional advantage that it is the primary material itself which is being consulted, avoiding, to an extent, a problem which was mentioned by one of those interviewed, of the secondary material distorting the primary literature. [MRC/SSRC 3B 210].

Library Catalogues, Abstracts and Indexes

An obvious method of starting in a new area is to consult bibliographies, abstracts, indexes, and library subject catalogues. This way of starting is that which would be most commonly understood as constituting a conventional literature search, despite the fact that a frequent observation of studies of the information seeking activities of academic social scientists is the relatively low importance attached to this means of locating information (Bath University, 1972: Hogeweg-De-Hart, 1984: Line, 1971).

Most of the social scientists interviewed had, at some time, undertaken a formal literature search of this kind, often when first embarking on their research careers. Some considered that if they were to start in a new area they would carry out some form of literature search through library catalogues, subject indexes, bibliographies or abstracting services [Psychology 7A 026-057: Continuing Education 9D 236-251: Economic and Social History 5C 487-491: Geography 7C 340-407: Sociology 13C 091-128; 11D 250-257]. But the use of such services was not heavy, and one of those interviewed expressed the view, which might be implicit

in others' non-use of abstracting services, that the usefulness of such services is rather restricted. [MRC/SSRC 7A 2655-2667].

Online searching of databases offers a quick alternative to manual searching of secondary services. All of the social scientists interviewed were asked if they had ever had an online search undertaken for them. Their opinions on the utility of these searches was mixed. In general there seemed to be greater satisfaction with the results of the search when the person 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 093 + 122-140: Continuing Education 10D 170-183: Education 2D 288-300; 1D 112].

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, or lack of suitable databases, it was clear that the online services were not seen as a sophisticated alternative to the social scientist's other information seeking activities. [MRC/SSRC 2D 219: Education 4D 407-454; 5D 290; 8D 296: Sociology 11D 161-193].

The comments of the social scientists concerning the problems with terminology, which are similar to those noted in a study of educationalists searching the ERIC database (Hounsell, et al., 1977), highlight the reservations concerning the reliance on terminology of both the information retrieval model and the cognitive approach. It is, however, worth noting, that even where the results of the search itself were not thought to be satisfactory the search 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 person could proceed to follow other information seeking activities such as following chains of references. [Education 6D 500: Sociology 11D 369-410].

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. In fact, often the social scientist starting out on a new topic is looking for something which can serve as a starting point for chaining. [MRC/SSRC 12A 2248-2249; 5D 237 + 471-480; 7A 2572-2593; 3B 175: Psychology 2A 545-657: Economics 1C 2298-2315: Prehistory and Archaeology 15D 068-078: Sociology 11D 198 + 390; 14C 064-072].

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 (1984), who has also considered the need for a theory of citing (Cronin, 1981). 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 shall be referred to here as chaining. Chaining can take two forms -

1. backward chaining - following up references or sources cited in material consulted;

2. or 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 academic 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 (Line, 1971: Skelton, 1971: Stenstrom, and McBride, 1979). Forward chaining, except in the context of legal citation, is a relatively new innovation, and dependent on the use of special bibliographic tools.

Backward Chaining

Following up references or footnotes was a major characteristic of the social scientists' information seeking patterns. All of the social scientists interviewed made some mention of it, and many employed it as their principal means of gathering information. [MRC/ESRC 3B 175: Continuing Education 10D 144; 11C 041: Education 4D 144: Economics 1C 216; 3D 057: Prehistory and Archaeology 15D 068-078: Sociology 10C 421; 12C 096; 14C 131]. 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 372-476: Psychology 2A 545; 9A 2077; Politics 9C 238].

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. Since 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 318: Economic and Social History 5C 109-119: Politics 9C 190-204].

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 565]. 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 person's notice through his other information seeking activities. [Psychology 1A 642].

Line (1971) had noted that over half of the

respondents to the INFROSS questionnaire had indicated that a citation index in the social sciences would be very useful, and Garfield (1963: 1964) had considered that citation indexes might prove more popular in the social sciences than in the sciences because of the difficulties created for conventional indexes by the changing terminology of the social sciences (Garfield, 1963: Garfield, 1964). 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 2049].

Although many of the social scientists were either unaware of the existence of these indexes or of what could be done with them, many 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 social scientists. [MRC/SSRC 9B 322-341; 10A 503-549: Psychology 4A 503: Economics 1C 348-403: Prehistory and Archaeology 15D 210: Sociology 13C 493 + 671].

Closure

The making and chasing of citations involves a considerable degree of subjectivity, and citations may be made for a variety of reasons (Cronin, 1981: 1984: Martyn, 1975: Moravcsik and Murugesan, 1975: Weinstock, 1971). Nor is it always the case that citations are considered important, or even particularly relevant by either the citing authors or the authors cited. Prabha (1983) in a study in the area of business administration found that less than a third of citations were considered essential by those citing. This phenomenon has already been noted in connection with the author assessments of their cited references to the search questions (Cleverdon, et al., 1966), and was remarked on by a number of those interviewed. [MRC/SSRC 4A 491 + 2217: Psychology 6A 2266; 2A 375-401].

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 220 + 656: Psychology 6A 2266-2297: Economics 1C 348-403]. Decisions to follow up particular trails are related to the

interests of the person 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 2232-2352: Psychology 4A 213; 3A 107: Education 5D 326: Prehistory and Archaeology 15D 087].

The extent to which an individual will wish to follow up references in material is affected by a number of factors. To a certain extent chaining may be intensified at the initial exploratory stage when a person 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 that person gathers information. [MRC/SSRC 1A 2038-2055: Psychology 9A 2105: Sociology 10C 427-428].

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 person has built up, or are becoming more and more marginal. The citations may also be leading off into areas of literature that the person 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 2065-2075 + 2173; 2D 169-174: Psychology 9A 2116: Sociology 11D 390].

Browsing

The concept of browsing has several different connotations, and a variety of different types of activities have been associated with it Ayris (1985). Nevertheless, whatever connotation is employed browsing is a recognised, if sometimes disparaged (Urquhart, 1976a: 1976b), form of information seeking. A number of different typologies of browsing have been put forward (Apted, 1971: Celoria, 1968: Herner, 1970: Levine, 1969), and a various studies have been made of the

importance of browsing, particularly in relation to the the question of the value of providing direct access to library material (Greene, 1977: Hyman, 1972: 1982: Lawrence and Oja, 1980: Ross, 1983). If purely random browsing is excepted, 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.

Semi-Directed or Semi-Structured Searching

Browsing, in the sense of semi-directed or semi-structured searching, was an activity that many of the social scientists 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

138: Economic and Social History 5C 089-106: Economics 1C 258-281 + 321].

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 browsing, in the sense of semi-directed or semi-structured searching, 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 235-244: Education 8D 029-034; 7C 200-215: Geography 7C 146-158: Sociology 12C 124-140]. 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 287; 4A 574: Education 3C 150-173].

Abstracting services may be treated in a similar way. The searcher browsing through issues of the abstracts not particularly looking for a particular item or for material on a particular, well defined, topic, but simply to check on the kind of work being carried out in an area or to see if there is anything interesting which might be worth reading or following up. [Education 3C 562: Economics 2D 591]. 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 415-419: Sociology 14C 127 + 319-333 + 486-490]

In addition to its role in the identification of material, browsing can also serve the purpose, 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 572-595: Education 5D 159-170 + 199-212: Economics 1C 258-292].

Differentiating

Brittain (1970) observed that in the course of the American Psychological Association project (1963: 1965: 1969) the initial impression that the communication system was characterised by confusion was displaced by the realisation that there were significant underlying regularities relating to the type and timing of publication of different types of information for different kinds of users.

> 'Contrary to the first impressions gained by some of the staff on the American Psychological Association project, the results showed that the system exhibited impressive regularities. Information flows through the system in an orderly manner and, although there are various routes, specific kinds of information produced by specific types of researchers seek certain outlets on predictable occasions in predictable sequences and time patterns. The outlets chosen by the researcher are very often associated with the specific needs of the user, and the information is shaped and reshaped to fit the characteristics of the channels and the needs of the users' (Brittain, 1970:97)

Garvey and Griffith (1972) 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. Brittain (1970) also noted that within social science disciplines there were often fairly well defined author and journal hierarchies. Brittain considered a number of studies which ranked journals in different fields according to the frequency with which they are cited (Boll, 1952: Gerould and Warman, 1954: Xhignesse and Osgood, 1967), according to perceptions of their status, orientation, or quality (Jakobovits and Osgood, 1967: Shepard, 1962a: 1962b), or based on the rejection rates for publication (Lin and Nelson, 1969). He concluded that

> 'Although the communication networks may appear chaotic from the outside, to the experienced researcher there are a number of structures in the system that guide his activities. 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' (Brittain, 1970:138)

In a later review Brittain (1979) commented on the importance of different theoretical or methodological

'schools' to the social sciences and the effect this has on publication and citation patterns within and between disciplines, and on the information seeking activities of social scientists. For the purposes of this study this activity of discriminating between different sources, or types of source, of information by academic social scientists will be referred to as differentiating.

Differentiating is effected by the researcher identifying different sets of sources in terms of the differing probability of their having useful material. The manner in which the social scientists differentiate between sources of information, rather than how they differentiate the subject itself, is analysed here. However, as differentiating between information sources was used extensively by the social scientists interviewed to describe 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 perceived as having the highest likelihood of containing material which was relevant, at an

appropriate level, and of the right type.

The criteria which seemed most significant for differentiating material were

1. the substantive topic of study;
2. the approach or perspective adopted;
3. and the quality or level of treatment.

These differentiating features can be separated for purposes of analysis, however, in practice, they are often employed together. Moreover, the same criterion can be used in different ways. For example, the level of treatment criterion was employed by some 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.

It should be noted that, in the case of some of those interviewed differentiating a limited number of sources was not possible or was very difficult. The nature of the 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 580-589; Psychology 9A 689: Economics 1C 182-195 + 222-231: Economic and Social History 5C 084-105]

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 523-544; 2D 109-124; 8A 166-178 + 329: Psychology 8A 091-094; 2A 131-147; 4A 052-065: Continuing Education 10D 030-043: Education 5D 159-170 + 199-212; 6D 100-141: Economics 2D 626-655: Geography 6C 031-050: Politics 8C 165-174].

Differentiating by Approach or Perspective

Differences between different 'schools' of thought in the social sciences have been studied in the context of a number of different subjects (Brittain, 1979: Kirsch, 1977: Krantz, 1965: 1972: Russett, 1970: Samelson, 1978: Weimar and Palermo, 1973). 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 at different levels ranging from the disciplinary approach to the topic, to differences of conceptual interpretation or methodology. [MRC/SSRC 2D 074-081 + 106 + 224-228; 9B 591: Psychology 7A 097-123: Education 2D 124-151; 3C 146-173].

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 096-123: Prehistory and Archaeology 15D 177-204].

Differentiating by Quality, Level, or Type of Treatment

As already mentioned, 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 325-360 + 435; 4A 566; 5D 202-215: Psychology 2A 2145-2154].

Material which is apparently on the right topic or in the right area may be treated at a level which makes that material inappropriate for the person carrying out the search. This may be because the material, though on the right topic, is too technical or completely incomprehensible to the searcher. Alternatively, the material may be set at too low a level for the needs of the searcher. In connection with the type of treatment, an important distinction to note is that made 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 person with a useful filter on the level and type of material examined. [MRC/SSRC 4A 016-028: Continuing Education 9D 037-072: Education 7C 560-576; 8D 057-071 + 535-543: Economic and Social History 5C 109-119 + 280-297; 4C 062-083 + 139-158: Politics 9C 190-204 : Sociology 13C 537 574; 10C 059-069].

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. Principal ways in which the social scientists 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 441-454]. and a person who has been working in an area a considerable time will be familiar with the existing sources in the area. [Geography 6C 093-104]. 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 375].

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 come up is correspondingly less. [MRC/SSRC 1A 490]. 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 person may feel they have to keep up with in more general terms - say for teaching purposes. [Sociology 10C 146 + 152-155]. 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 166-178].

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 318: Continuing Education 11C 186-287: Education 4D 330-338]. 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 211-225: Sociology 14C 127 + 143 + 146].

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' (Price, 1961: Price and

Beaver, 1966), and that of the 'gatekeeper' (Allen, 1965: 1977: Allen and Cohen, 1969). 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 (1982).

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 informal contacts. 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 589; 5D 393-413]. 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 040-048 + 070-081 + 098-110].

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. In this way social scientists 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 6A 2320; 1A 689:

Politics 9C 515-540: Sociology 10C 094-101 + 439-445].

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, the person may rely on these individuals or organisations to filter material for him. [Education 3C 188-197: Politics 8C 061 + 088-114 + 555: Sociology 10C 112-115].

Monitoring Journals

Skelton (1971) 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. Ford (1977) found that only 17% of social scientists considered journals to be their most important source. 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 frequently publish material of interest. 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; and indirect awareness of the existence of other sources, and of material in them, from references in the sources directly monitored. [MRC/SSRC 12A 278 + 344: Psychology 2A 203-216: Continuing Education 11C 186-267; 9D 037-072 + 121: Education 5D 159-170; 7C 200 + 222-233: Economics 1C 321: Geography 6C 031: Politics 8C 232 + 240: Sociology 13C 167-187]. 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 146-173 + 359-360: Economic and Social History 4C 097-103: Politics 9C 059-067 + 190-213].

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 2604-2607; 10A 252-282; 4A 234: Psychology 6A 580; 4A 026-046 + 075-078; 5A 315 + 480; 7A 068 + 097 + 433-470; 8A 098 + 296-297; 9A 440 + 452; 1A 206: Education DAT 092-123: Economics 1C 202-214 + 258-281 + 292].

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 146: Continuing Education 10D 022-038: Education 2D 114-122; 6D 215: Economic and social History 4C 168; 5C 234; Geography 6C 031-050: Sociology 12C 096-112].

Extracting

Extracting refers to the activity of going through a particular source selectively extracting material from that source. The source may consist of 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 person would attempt to make a direct examination of the sources themselves. [Psychology 2A 203-218: Continuing Education 9D 162-181: Prehistory and Archaeology 15D 243-262].

The identification of potentially useful sources of material is a critical pre-requisite 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 3024; 8A 166-178 + 329] 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 123-125]. Sources may also be identified through browsing. [Education 8D 027-039].

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.

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 2077: Education 8D 027-039; 6D 500-518: Continuing Education 9D 181: Economic and Social History 4C 456-473: Economics 3D 038-043: Geography 6C 031; 7C 335: Politics 9C 240-254: Prehistory and Archaeology 15D 632-634].

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 135-149: Economics 3D 038-043: Geography 7C 098: Politics 8C 240]. 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 a person's interests may become relevant or important not only if the person'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 109-124: Psychology 4A 052-065].

Extracting from Publishers' Catalogues

The use of publishers' catalogues for monitoring publications in book form has its complement in extracting. For many of the social scientists 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 pre-selecting those catalogues which they wished to see. Others had made direct contact with the publishers or bookshops. [Education 7C 259-264; 1D 041 + 046; 2D 108-112: Economic and Social History 5C 234: Economics 2D 161: Politics 8C 190-195: Sociology 10C 083-086; 11D 220].

Extracting from Bibliographies, Indexes, and Abstracts

Although not representing a major aspect of the information seeking activities of the social scientists 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 088 + 225-237: Psychology 6A 2511-2520: Prehistory and Archaeology 15D 640-692: Politics 8C 190-195].

Conclusion: the Nature of the Behavioural Model

The behavioural model outlined here may be described in a number of different ways. In the empirical derivation of the features of the model from the empirical data, and its development in relation to a substantive area of social inquiry, it has many of the characteristics of Glaser and Strauss's (1967) description of a grounded substantive theory. The construction of the model from a number of semi-independent components corresponds to Diesing's (1971) description (derived from Kaplan (1964)) of the concatenated structure of holistic theories. In

Diesing's (1971) account the behavioural model would be described as constituting an empirically derived typology. The behavioural model can also be understood as representing what Paton (1980) describes as an analyst-constructed typology.

In form, therefore, the behavioural model represents a relatively conventional intellectual construction. However, the model is not primarily intended to constitute a means of describing or explaining the information seeking activities of social scientists, although, of course, it may provide some insight into those activities. Instead, the intention is to employ the model to derive recommendations for the design of information retrieval systems for academic social scientists.

In the following chapters the features of the behavioural model are set out with reference to the details of the information seeking activities of the various groups of social scientists interviewed, this is both to illustrate the nature of the evidence on which the model is based, and to show that the framework of the model is sufficiently robust to serve as the basis for the design of information retrieval systems for academic social scientists.

Chapter Four

MRC/SSRC Social and Applied Psychology Unit

The MRC/ESRC Social and Applied Psychology Unit is separate research centre of the University, attached to the Department of Psychology. The staff are engaged on full-time research projects for the Unit's sponsors, and on other contracted research projects with interested organisations. There are ten permanent staff in the Unit, and, at any one time, up to fifteen other research staff working on the various projects. At the time of the interviews the Unit's principal research

areas were in occupational, clinical, social and cognitive psychology. The name of the Unit at the time of the interviews was the MRC/SSRC Social and Applied Psychology Unit, therefore, for historical accuracy, it is that title which is employed here.

Starting

Several of the researchers mentioned that it was unusual to be faced with starting out in a completely new or unfamiliar area and that often they were able to begin from what they were already aware of in the form of starter references. The extent to which the researchers provided for this varied between that of the researcher who maintained files of potential starter references and that of the person who simply made a note of things which he thought might be of use in the future.

'These things rarely happen like a bolt out of the blue. You don't just suddenly get presented with a new topic. Over the years you build up expertise in certain areas but then you do know a little bit about other areas, and you would know where to start looking. You would probably know a few key people or key references and you would go to them and build up - network like. When I think about certain topic areas that have developed over the years that's what's happened...I have these working files which I shove things in. They may be papers, or notes, or ideas I've had for analysis of some existing data. I've not had time for them so I've put them on one side under a topic heading. So it will be at an embryo stage - it won't be just a topic heading on a piece of paper you've always got something to go on' [MRC/SSRC SAPU 1A 650]

'Others I might happen to come across just by accident and think that may be useful one day. I may want to know about that. Not something I was looking for at the time - but something I think I might need that one day...Putting it away because I think I might need it. To some extent it is hoarding' [MRC/SSRC SAPU 8A 250-260]

Research assistants and research students coming into the Unit were often provided with references by established members of the Unit. This way of inducting new researchers has the advantage of saving the newcomer time identifying pertinent references and allowing the selection to be based on the criteria of topic, importance and perspective which the other researchers in the Unit are operating with. In this sense, its value is not restricted to that of familiarising the new researcher with the material, but

extends to that of familiarising him with the ways of approaching the subject which the other researchers consider important.

> 'What we have done in the past, by and large, we give them all our own stuff to read first - so they get some idea - because apart from anything else that probably refers to 200-300 publications over about four or five papers that we have done recently' [MRC/SSRC SAPU 12A 577]

> 'When they come in here they are going to spend their first three months pretty well with their head down and immersing themselves in the literature. And I can save them a lot of trouble there and so can my colleagues. We can say here's an index file of everything that we have found on this topic - but actually these are the ones you need to read - and pull out half a dozen of the most important ones. And, when they've gone through those say well you can now have a look at some of these other ones' [MRC/SSRC SAPU 7A 694]

Many of the researchers mentioned reviews, and review type material, as being particularly valuable when starting. Not only were they considered a useful source of references to other material, but, also, by their very nature,they provide an overview of the topic. Reviews could therefore provide references to other material, synopses of work done, introductions to key concepts and key workers, and a means for the researcher to start to organise his own thinking on a new topic.

'When you first decide you want to get into a new field you look at review articles in the area to get an idea of what the main concepts are, and who are the main people, and what's been done up to now' [MRC/SSRC SAPU 8A 364-371]

'One will be good thorough review papers. You'll grab those because they'll help you when you come to do your own reviews and organise your thoughts, and provide you with a lot of references that you can chase up if you wish to. Any good review I come across in my area I'll keep' [8A MRC/SSRC SAPU 275]

'I know within a matter of five minutes I could find review articles, I could find bibliographies, that would give me nearly every reference I would need anyway...so I have reviews and we have our own articles and we have compendiums which we can get hold of' [MRC/SSRC 5D 490-510]

'Perhaps initially I would direct somebody towards the conceptual or the theoretical writings...and then there's also a set of empirical studies...to see what has been done and what we know about this field. There is a third group of studies which perhaps you'd put the person in touch with first of all, and these are the review type articles and summaries of the field' [MRC/SSRC SAPU 7A 2011-2027]

The form the reviews took and the manner in which they were identified varied. Handbooks and yearbooks were mentioned as potential sources for reviews of particular topics, or of recent work in an area. Some

journals provide annual reviews of work carried out, others are recognised as being more likely than others to publish review type material. Into this category, also, come journals which publish synoptic tutorial articles on particular topics which can serve as reviews.

> 'There are handbooks and yearbooks, substantial tomes which give an overview of a particular field. That's the kind of reference you would give more prominence to if you wanted to help someone gain an overview, and of course reviews. Things that aren't simply one study empirical reports - these may be very good in their own terms - but not in terms of helping someone understand the whole field because they may be very specialised and narrow' [MRC/SSRC SAPU 1A 2145]

> 'The Psychological Bulletin, for instance, is a journal which will have a disproportionately high number of theoretical review type studies. There's also things like the Handbook of Organizational Psychology - the chapters of which will give an authoritative overview of research in a given topic area. They are the most useful of all - these are the reviews done by the academics' [MRC/SSRC SAPU 12A 2400-2414]

> 'Computer Surveys is bloody good. I use that. The long article about text editing was in Computer Surveys - there are a very large number of references in that...These synoptic articles are written by people who are leading experts in the area, and the publication lag is not too long. By the time you've read them they are not too out of date. It's also written deliberately as a tutorial journal. Explaining an area to

> people who are not necessarily intimately acquainted with it yet at the same time trying to get up to the state of the art. Hard work writing for it I should think' [MRC/SSRC SAPU 4A 3122-3131]

The reliance on review type material for starting in an area is not without disadvantages. The reviews may present a view of the area which distorts the information contained in the primary material. This point was made strongly by one of the researchers interviewed.

> 'I feel there's a stereotypical view of the impact of unemployment on people, which to a great extent is the result of secondary sources streamlining the primary material...The secondary literature distorts by omission' [MRC/SSRC SAPU 3B 210]

The use of collections of primary material may partly avoid this problem, though any criterion of selection employed might be considered to be subject to some bias. Nevertheless, such collections may represent a useful way of starting in an area providing both an overview of work done, and a useful set of references.

> 'The easiest thing to pick up a volume of collected papers...More collected papers than review articles... I've got one in mind here...Now this is a collection, but also quite up to date, and with an excellent introduction, and this, taken together would constitute an excellent review of the field. If someone wanted to get into this area very quickly I don't think I could suggest anything better than first read that. They're all previously published articles which have

> in some way been germinal in stimulating work in the field, and it gives a good topic coverage and very good papers...I think that its performed an excellent service by being there, because referring out from that, one would cover a great deal of literature' [MRC/SSRC SAPU 11A 539 2099-2169]

The use of secondary services, in the form of library catalogues, abstracts and indexes was not particularly remarked on by the researchers interviewed. This may have been partly because of the nature of the work of the Unit, and one researcher specifically commented on the limited value of abstracting services for project based research.

> 'I think that you will find that in most cases this is true that the abstracting services are not designed for the idiosynchrasies of given research projects. They are design in order to divide the field up in a certain way. I think that they are probably useful for PhD students at the first stage. But once you've got into - once you've identified a few key things, you will soon get past the stage when abstracting is very much help to you' [MRC/SSRC SAPU 7A 2655-2667]

Some of the researchers had had computer based searches undertaken for them by the University Library. These seemed to have been undertaken when the researchers concerned were moving into an area with which they were unfamiliar and were looking for a quick way in, or where they perceived that material might be widely scattered. In either case, the intention seemed to be to use the computer search to provide a starting

point into the literature, rather than to constitute a comprehensive bibliography complete in itself.

> 'If it's a new area you might just specify four or five keywords...The Library here will do a computer search for you using keywords and that is just a general one shot - to see what's going on to see what the new field's like' [MRC/SSRC SAPU 5D 093]

> 'Something like mental health...there are fewer studies and we know that, and...it's one of those areas where the sort of articles we'd be interested in might be in different sorts of journals, which we might not...either know of or know the authors, so it's a search in that sense to try and find out what's been happening in areas we don't really know about and in journals we don't know about' [MRC/SSRC 5D 122-140]

It is interesting to note that the limitations of such a search may be recognised, and, to a certain extent, compensated for, by the searcher applying more sophisticated cognitive criteria to the resulting output.

> 'You've got to filter them because it just works off the keywords - as I understand it - from what I remember we got some that were spot on and clearly important and some that weren't. And, of course, typically, if you're looking at it from a psychologist's point of view, and a psychological focus then you will find these in psychological type journals - which makes the searching easier' [MRC/SSRC SAPU 2D 219]

The relationship between starting and chaining was very close. Starting would very often change almost immediately into chaining as the researcher identified a few references from which he could proceed to backward chain, and chaining would often start from a few starter references, or from a review, or from the result of a computer based search.

> 'You've some starter references and in you go - and you'll start from the most recent backwards because they obviously can refer back and the forward ones can't refer forwards' [MRC/SSRC SAPU 12A 2248-2249]

> 'I moved into the shoes of someone who was leaving who had been doing some work on unemployment so I took over their pile of offprints, then I really worked from references at the back of each article, getting hold of the articles that were referred to that seemed to be at all relevant and then using their reference lists. that very soon multiplied into a large number of articles. [MRC/SSRC SAPU 3B 175]

> 'Try and get hold of most recent reviews of a particular area and follow up the references given in that particular review' [MRC/SSRC SAPU 5D 237]

> 'I would use some of the references taken from the computer search and I would look through those articles to see if there are any references contained within those that might lead me off in another direction. It might be away from the computer search into a

> different sort of field which I hadn't thought of until I read the article that was brought up by the computer search - so that we are on different levels of analysis. A computer search can be a very broad sweep of an area or a very specific dip into an area, but I would still look through the articles that it threw up and see if that led me anywhere else anyway' [MRC/SSRC SAPU 5D 471-480]

This way of proceeding was perceived as representing a reasonably reliable means of becoming familiar with the literature of an area, and one researcher commented that this method was more efficient and effective than reliance on a 'keyword' type of search undertaken by an intermediary because it allowed the criterion of relevance to develop as the search progressed.

> 'I would start out by finding some of the recent key articles. I suppose you latch on to, you know who the names are in a field, and you look at the books or the articles that they've written, and then from those you then chase up their citations basically, and that's the way I would develop actually, I would work through the literature in a sort of branching sequence of reading things and seeing. Which is slightly more systematic than getting someone else to do it by critical work keywords, and so on, because on the way you develop criteria of relevance or half relevance' [7A MRC/SSRC SAPU 2572-2593]

Chaining

Backward chaining was mentioned by all the researchers interviewed. Often the activity was iterative and cumulative with references from one article being used as sources of further references.

> 'I really worked from the references at the back of each article. Getting hold of articles that were referred to that seemed to be at all relevant, and then using their reference lists - that very soon multiplied into a large number of articles' [MRC/SSRC SAPU 3B 175]

> 'But then you chase them up, you find one article that has got references to ten others, so the thing escalates' [MRC/SSRC SAPU 12A 372-376]

Forward chaining using the Institute for Scientific Information's Citation Indexes, although not employed as generally, or as often, as backward chaining, was mentioned by several of the researchers in the Unit. The choice between Science Citation Index and Social Science Citation Index depending on the area in which the researcher was working. Forward chaining was perceived as representing a useful way of locating

additional references, particularly where the researcher knew of a key paper or author, or where it was desired to chart the impact or influence of a particular author or piece of work.

> 'The score at the end was that Science Citation Index/Social Science Citation Index worked quite well. I got hold of a *fair* amount of material from that' [MRC/SSRC SAPU 4A 2049]

> 'Forward chaining and backward chaining - but only the ones that looked interesting. Forward chaining using the Science Citation Index - Social Science Citation Index...went through a whole sequence of one guy's work - a sequence of very similar experiments - but I wouldn't say I use them very often' [MRC/SSRC SAPU 9B 322-341]

> 'I have sometimes used the citation index - the Social Science Citation Index - if I'm wanting to update on a field that I can identify with a particular paper...I've found that more useful I think mainly because of the facility of following a net - following by a sort of branching thing from known papers, I think that's the main reason' [MRC/SSRC SAPU 10A 503-549]

One researcher, with reference to citations to a piece of his work, and to citations in his field generally, noted that material may be cited which has only an indirect connection to the subject of the article. This may be because the author wishes to indicate familiarity with material, which, in fact, has

little relation to the citing work, or may be citing others work for reasons which are largely social or professional.

> 'There is a hard core of papers which are referred to over and over again. The paper we published in 1975 is an example. Its very frequently cited, cited much too often in fact in cases where it just is not relevant to show that people have seen it...If it comes from certain places, particularly in the USA, then it's got a better chance of being cited very often than from other places. That's how it goes. They've always got a lot of their students out there writing papers, and they are going to cite the papers their professors write aren't they? Fair enough so would I' [MRC/SSRC 4A 491 2217]

The extent to which the researchers followed up references was also variable. As part of the function of citation is to indicate connections with other related material a large element of cross referencing is to be expected. Researchers who had been working in an area a considerable time were, naturally, often familiar with a large number of the references made. The extent to which the researcher would be familiar with references in a particular article or book, largely depended on the centrality of it to his own work. If the topic was highly germane to his work then a large proportion of the references might be known. If other topics were covered which the researcher was not so familiar, then the references to that topic were not so likely to be known.

'One doesn't have to go too widely in the standard journals to get cross referencing - a lot of other people do a lot of work for you. You're reading something - you say o.k. where did he get his stimulus. If he works from a literature stimulus where did it come from - because these are on his reference list - o.k. and one can branch out from there and follow what one feels like following' [MRC/SSRC SAPU 11A 220-253]

'It's very rarely that any paper is a blazing novel experience. Generally speaking papers are building on prevailing thinking...Were it a paper on language syntax - comprehensibility of language syntax - then I should be disappointed if I hadn't seen 90% of the references, but if it was something a bit further out, in a field that I'm not into, such as memory models or something it may be that a higher percentage of recent works references I wouldn't be familiar with' [MRC/SSRC SAPU 11A 656-679]

'I always look at the references quite closely...and I'd say that I know or have already on file somewhere or other about 85% - 90% of them...the ones that aren't will either be on specific points raised which are probably tangential to the topic. For example if X is writing on stress she might have a lot of references on stress. If she's writing on stress and careers - which is the kind of thing she would write on - the references on stress I probably wouldn't be familiar with - but then I wouldn't be very interested in them either. But within the area in which I'm working I'd know them all, and if there was one I didn't know I would look a it fairly closely to see why I hadn't found it. I pick up maybe - each time one of these people writes an article I pick up maybe one or two references that I might have missed, but even that's quite unusual really' [MRC/SSRC SAPU 7A 2252-2352]

The comment of one researcher indicates the extent to which personal, and, especially, time factors, can influence decisions as to how far material is followed up.

> 'It's always task focussed. You've got to give a paper or you're writing a paper and you need some information or support. Then you start following things up and you can't let them lie...If it's a literature review you may have to trace all those under the topic. It's quite rare. Most papers are introduced quite perfunctorily with a few key references - and thats fine. There are no real rules. You say do I really need that paper or not' [MRC/SSRC SAPU 1A 2038-2055]

Nevertheless, there are signs which provide an indication to the searcher that the search has been reasonably comprehensive and that most of the pertinent material has been identified. For example, where new material is referring to material already known, and less and less material is being cited which is not known.

> 'You know that you've reached the end of your literature search when your starting to pick up things that are referring to things that you've already heard of...They obviously refer to a few things that you haven't heard of for a very good reason, they've used a particular measure that we haven't used and they've got to give references to that. And, when you get to that point you begin to feel fairly confident that there can't be that much else around. Otherwise you would have heard about it. So there's a sense of closure I suppose when you've completed your search (although you're always aware there must be something you haven't heard of, or something that you could refer to at hand)...The

subjective critical mass of known references' [MRC/SSRC SAPU 12A 2065-2075 2173]

'I find, and I'm sure that others do, that once you're into it you keep on - you know you're getting somewhere when you keep on seeing the same things being cited again and again as important works and after a period of time you're finding less and less new material' [MRC/SSRC SAPU 2D 169-174]

Browsing

The researchers interviewed did not make much mention of browsing. Where browsing was mentioned it was in the context of browsing through the contents pages of journals in their areas. This may be connected to the use of Current Contents - which is treated here as an aspect of monitoring - and which may represent a substitute for actually browsing through journals.

'You don't browse through everything! I do browse through journals hoping I'll pick up something...There are some journals which might have articles in which I'm missing. You make some sort of educated guess on the basis of past history of interesting things that you've met as to where the interesting things are most likely to crop up in the future' [MRC/SSRC SAPU 9B 287]

'All the journals that are remotely relevant in the library here I skim' [MRC/SSRC SAPU 4A 574]

'You pick up a journal on human computer interaction and look at what people have done, and if it says something like 'the effects of consistency and and compatibility on command language learning' then I automatically think - ah that looks interesting, must read that. That's a very specific example - more often I browse and something hits me as being interesting. I'd read the title first - I think probably on the basis of the title I'd look at the abstract. I wouldn't read an abstract from a title I was certain wasn't going to be interesting - obviously. Some titles give less away than others' [MRC/SSRC SAPU 9B 235-244]

Differentiating

Differentiating between sources of material was a key aspect of the information seeking activities of many of the researchers interviewed. This operated as a filter on the number of sources directly examined. The basis for differentiating between sources was the differing probability of the sources containing material of interest. To effect this three principal differentiating criteria were employed, differentiating by substantive topic, by approach or perspective, and by level or quality of treatment.

Differentiating by substantive topic is an obvious criterion, but it should be borne in mind that the differentiating criterion here is being applied not merely to the subject but also to the sources. Furthermore, application of the criterion is not limited to the researcher identifying material on a topic, but extends to the identification of sources which specialise in, or regularly carry material on, the topic.

> 'I always look at the epidemiology, and social medicine, and health education journals...In Current Contents I look at all the preventative medicine journals. The

> occupational psychology ones - there are fewer anyway - so I tend to look at all of them' [MRC/SSRC SAPU 1A 523-544]

The contribution of differentiating sources in terms of their substantive topic to the researcher's overall information seeking pattern was particularly apparent where the researcher had changed interests, and, as a result had started to consult entirely different sources.

> 'You start looking at different journals. Since I've been interested in health I've looked at the Journal of Health and Social Behaviour and a journal called Social Science and Medicine. A few years ago I would never have looked at those journals...As soon as I got the health component into my work it began to make me look at some new journals. But there have always been articles on stress at work and aspects of health work in the journals I've traditionally read. But once I'd got a specific interest I began branching out in that area. So I've begun to read the Journal of Stress and the Journal of Psychosomatic Medicine, which I wouldn't have done before, not look at those as journals. I might have picked up articles in them - even published in them - but I wouldn't systematically have read them. I'd have picked them up from reading the occupational journals...When you get into a new field - when I got into stress - I really didn't know about the Journal of Health and Social Behaviour, because it's a sociological journal, as is Social Science and Medicine. But when I became interested in health I came across references to it which I though were good - so I'd go and look more carefully at those journals. You begin to explore and then concentrate on those you like and they tend to lead you to others anyway' [MRC/SSRC SAPU 8A 166-178 329]

Sometimes it is the source which changes, and this,

rather than any change in the interests of the researcher may lead to the source becoming key.

> 'Something can become relevant either because it's new and therefore is coming out in an area that is relevant to the Unit or it can become relevant perhaps because there has been a slight change in editorial policy, or perhaps we have slightly changed areas and something that that wasn't relevant suddenly becomes relevant - or, you know, there is a sudden growth of interest in a particular journal and people start publishing in it about such and such a thing' [SAPU 2D 109-124]

Differentiating in terms of approach or perspective can take a number of forms. Differentiating may be in terms of the type of approach characteristic of a whole discipline, or refer more specifically to specialisations or perspectives within a discipline. Divisions of this nature are reflected in, and perhaps to a certain extent perpetuated by, divisions in the journal literature. The division of material from different approaches or different specialisms into different groups or types of journals was remarked on by a number of the researchers interviewed.

> 'These are the sorts of things that people in the area publish in and use and refer to...and, of course, typically if you're looking at it from a psychologist's point of view, and a psychological focus, then you'll find this in psychological type journals, which makes the searching easier' [MRC/SSRC SAPU 2D 106 224-228]

'I am not a cognitive psychologist but (colleagues) have a different perspective and read other things as well and talk to other people. They read different sorts of journals, and they write different sorts of books, and they write different articles because they have a different focus on it' [MRC/SSRC SAPU 2D 074-081]

'Very little of linguistics would be of direct interest to me. Most of computational linguistics wouldn't be of interest to me at all. Or the interest would be quite irrelevant to what I was doing...I can't understand them anyway! They don't appear in the same sort of journals. I don't look very often at the Journal of Computational Linguistics. I don't think I understand it when I do. I have a notion of what's going on in that field, and I vaguely understand bits of it, I couldn't understand a detailed technical paper about it' [MRC/SSRC SAPU 9B 591]

'Part of it is simply that, whatever the notes for contributors say, a particular journal will tend to have a theme or themes. What happens is that papers that cite certain other papers tend to be located within that particular journal, rather than some other journal. Even though in terms of the notes to contributors they might formally be just as appropriately placed there...You get a tendency for a particular journal to repeatedly publish from people in this little group of people who are citing each other and are developing their own little sub-field. So that's a criterion, as it were, a criterion of that's the sort of journals where the people and this sort of stuff tends to come out' [MRC/SSRC SAPU 10A 325-360 435]

One researcher commented on the fact that, in terms of the actual subject matter, there ought to be

more of a connection between different aspects of the topic than there actually was in the literature. There was a strong tendency for material on the different aspects to appear in different sets of journals, which were themselves associated with the different professional groups dealing with the subject.

> 'It's interesting because conceptually, theoretically, there ought to be more connections than there are between the study of the things that bring people to have psychological problems on the one hand, and the mechanisms that are thereby revealed and the process of treatment and how you help people overcome these problems on the other. There ought to be a much closer link than there actually is...There are some psychiatric journals that have both, like the Archive of General Psychiatry, this is an American journal, a very good journal, that has a bit of both in it...The psychiatric journals tend to have a bit of both in, I think that's a fair generalisation. The psychological journals tend to have more therapy research, and less of the life events and epidemiological research. The generalisation isn't absolutely watertight, but that's the general drift of it. And there are some journals that are specifically dedicated to treatment, and they have 'therapy' in their title. So you have Behaviour Research and Therapy, you have Cognitive Therapy and Research, you have Journal of Behavioural Therapy and Psychiatry. You have various journals like that which are clearly therapy journals...There are journals which have nothing - virtually nothing - on therapy in them, which will have life events and other things in them, like, for example, Journal of Health and Social Behaviour, or Social Science and Medicine - that won't have much therapy in it. So, if I'm reading in what we are calling life events research, then these journals will be very high on the list, and they will have a lot of that kind of work but virtually nothing on therapy' [MRC/SSRC SAPU 10A 096-123]

Differentiating by quality or level of treatment partly relates to straightforward judgements of different sources being likely to publish material of a different standard. If something appeared in a journal with a high reputation then the researchers would be more likely to feel that they should read it. Several of the researches commented that they would be less likely to feel that they should read material appearing in journals with not such a high reputation, because of the greater possibility that such material would not meet the appropriate standard.

> 'Usually you can tell on quality by the journal. On the whole you know what the reputation, the reviewing procedures, and the rejection rate are. Journals have reputations so you know that if something appears in a reputable journal its got to have got through certain screening procedures and it will probably be O.K. Some rubbish things do get through, but compared to some journals where they don't have a proper reviewing procedure - the editor might decide, or semi-popular journals - you have less expectation of them' [MRC/SSRC SAPU 4A 566]

> 'Rightly or wrongly, I informally operate the belief that there is a small number of top quality journals in which the relevant stuff - the best relevant stuff - appears, and therefore I don't need to go into those other journals in that detail...There's a notion of quality, there are good journals...practically all the good journals are American in both of these fields' [MRC/SSRC SAPU 10A 325-360 435]

'If you take straight line scientific journals, for example, then in our field it would be the Journal of Applied Psychology, Journal of Occupational Psychology, Journal of Occupational Behaviour, Academy of Management Journal and Review, and Human Relations, all of that sort. Then if it appears in these sorts of journals, which are refereed journals, then you know that there is some basis for why that article's in there. It has to reach some sort of scientific or professional criteria to be accepted for these journals, and you know that's been done by peer review, and you know that most of the articles in there will be of a reasonable standard. It's not saying all will, because some of them aren't, as everybody knows, but that's one criterion. If you know that it appeared in the Journal of Applied Psychology or the Journal of Occupational Psychology then there's something reasonable going on. If it appears in something like Personnel Management then it could be a good review type article in practical terms for managers, or it could be a one of those "well I happened to be in the factory one day" type of report - and that's more difficult to decide' [MRC/SSRC SAPU 5D 202-215]

Differentiating according to quality or level of treatment can also involve perceptions that different sources have different functions. So that the researcher would have different expectations of material in the different sources, and would use them for different purposes.

'The Artificial Intelligence community used to meet at conferences and not publish. It still meets at conferences and doesn't publish much. It now has its own journal called Artificial Intelligence. This publishes long papers and is consequently not right up to the minute. There is a long publication lag. That meets one of the main requirements for any science of being a respectable archive source for major

publications - that is all it does meet - it doesn't have any outlet worth speaking of for the shorter note. There is now something that is only semi-refereed and not archived. This serves the complementary purpose of publishing short notes of what is going on, what people are doing, what people hope to be doing, what the research application was, fairly informal accounts of conferences - who to write to for papers. It's right up to date. The current issue of Special Interest Groups in Artificial Intelligence has papers from a conference in January - so the publication lag is measured in months, you can find out what people are doing right now' [MRC/SSRC SAPU 4A 016-028]

Not all the researchers found that differentiating between sources to filter material was useful. Where the material was very scattered then identifying a limited number of core sources was not possible.

'It tends to be scattered I think. To be honest there's not really an awful lot of work that I'm very impressed with - it's the minority - and it can be scattered around...I get a lot of papers sent. They just come through the post or recommended. Our Director here monitors all the journals that we take and he will send photocopies through of anything he thinks might be of interest. Or else I'm reading one article - I'm struck by something in it- I look up the reference in the back and follow up that particular one. But there aren't any journals that I regularly scrutinise for content' [MRC/SSRC SAPU 3B 580-589].

Monitoring

Many of the researchers remarked that personal contact represented an important means by which they kept up-to-date with developments in their field. The MRC/SSRC Unit was particularly interesting in this respect in that the researchers were working in teams on projects and different members of the project teams could inform other members of their team, or colleagues in other teams, of work which they thought might be of interest. The different researchers were, to a large extent monitoring different sources, and so were able to keep each other up-to-date with pertinent material appearing in a wider variety of sources than one individual could monitor. In addition, the Unit's Director operated as a semi-formal gatekeeper for the project teams. He monitored all the publications coming into the centre and informed individuals of material which he thought might be of interest to them.

> 'Our director here monitors all the journals that we take and he'll send photocopiers through of anything he thinks might be of interest' [MRC/SSRC SAPU 3B 589]

'One of the best abstracting services the Unit has is the Unit Director - who finds things even before they come off the press. He will float pieces of paper around saying have you seen article so and so, and so and so. Which presumably takes up quite a bit of his time but is very nice from our point of view' [MRC/SSRC SAPU 5D 393]

Many of the researchers also mentioned monitoring journals as the principal means by which they kept up-to-date. Prior differentiation of a limited number of key sources allowed some of the researchers to concentrate their monitoring on a small number of journals, they would then rely on citations from articles in these journals to alert them to other pertinent material in sources not monitored directly.

'You read your journals as they come out...At this moment in time I will keep an eye on all the journals as they come out, and on the books as the publishers send them to me' [MRC/SSRC SAPU 12A 278 344]

'The Unit's got its own library so periodically I go over there and just wade through all the current journals, and most of those journals are the ones that we are interested in anyway...That's the way to keep up to date, it's just to keep walking over and having a flick through these or a flick through Current Contents' [MRC/SSRC SAPU 5D 393-413]

'I suppose I scan about six fairly systematically. I rely on scanning those to lead me to others. That doesn't sound very

> many but you don't need to - they lead you all over the place. One of the things you don't lack in this field is things to read - the problem is learning what not to read' [MRC/SSRC SAPU 8A 318]

In addition to actually monitoring the journals some made use of Current Contents. This way of monitoring journal contents, because it is less laborious than examining the contents pages of the journals themselves enabled the researchers to scan a very large number of journal contents pages generally, while allowing them to concentrate more attention on those journals which they had identified as being likely to carry material of interest. In some respects this might be seen as integrating certain aspects of browsing and monitoring.

> 'I look at Current Contents regularly, so that way I keep in touch with journals that we don't keep - we keep quite a few journals here you see so that makes it easier to keep in touch with the current literature' [MRC/SSRC SAPU 7A 2604-2607]

> 'I just use Current Contents. I know which journals on the whole I automatically look at. And then I have another second division group that I look at if I've got enough time' [MRC/SSRC SAPU 1A 490]

Current Contents was not, however, highly used or highly valued by all the researchers. One researcher remarked that he had employed it at one stage but that

his use of it had fallen off, another considered that the service was not particularly useful for his area of research.

> 'I go through phases - sometimes I'm an assiduous Current Contents person. But I haven't been that for a couple of years...Scan through I suppose what must be 30 journals...because there are psychological journals, there are psychiatric journals...For the last two or three years...that has slipped, and, in the life events research, I'm to some extent being fed by my colleague who will suggest to me papers, and so on. I occasionally make forays into the library...I know which of, in that field there are maybe only half a dozen journals that are going to throw up papers spot on what we are doing and I keep up with those' [MRC/SSRC SAPU 10A 252-282]

> 'I do have access to Current Contents but I don't use it very much because it tends to be full of - biomedical history of grass studies, plant psychology, a spoof area I made up! Training tiddlers to run mazes. In other words Current Contents is full of stuff which is not merely off beam to me but is actually bizarre. This is Current Contents in the behavioural area. Whereas in the hard sciences it stays more or less in the sane and rational, by the time it gets on to the behavioural sciences there are some very odd journals around. Any journal has papers that are odd by its standards - when you consider a paper which is odd by the standards of a journal which by my standards is odd!' [MRC/SSRC SAPU 4A 234]

Only one researcher mentioned making use of publishers' lists.

> 'I've got my name on book circulation

advertisements. Any new books that come out I get hold ot them. I read those lists fairly thoroughly' [MRC/SSRC SAPU 8A 146]

Extracting

Two researchers made specific reference to extracting from journals and abstracts.

'I...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 p ssible I would go through a few issues of that' [MRC/SSRC SAPU 4A 3024]

'Initially I would go through the back issues of some of the major journals, just to check on my knowledge of the literature in these major journals. I would then go to something like Psychological Abstracts, and if it had a more medical feel to it I would go to Index Medicus. I would also use Current Contents and go back through those. So a number of quick ways...Other ways are to try and get hold of most recent reviews of a particular area and follow up the references given within that particular review' [MRC/SSRC SAPU 5D 225-237]

'With international studies we rely on abstracting services, following up major journals, and contacts with people overseas, and the usual way you look through Psychological Abstracts or Index Medicus or whatever' [MRC/SSRC SAPU 5D 088]

Conclusion

The researchers in the MRC/SSRC Unit represent both the most cohesive and the most obviously research

oriented of the groups of social scientists studied. The researchers in the Unit are all working full-time on research projects. The different research projects can be undertaken without the distraction of teaching, and so the 'life-cycle' of the research projects is not distorted by the need to accommodate them to the demands of the teaching year. The researchers are not working in isolation from each other but with several others in a research group. There are close affinities of research interests both within and between the project groups, and the researchers are located in close physical proximity to each other. This enables efficient and effective cross fertilisation of ideas and information within and between the groups, which is also enhanced by the coordinating role of the Unit's Director, who, in addition, alerts the researchers in the Unit to material which might be of relevance to them. Finally, the nature of the research projects carried out, and the turnover of contract researchers means that there is a continous need to induct new researchers into the work of the Unit and to set them to work on new and existing projects.

It is clear that the researchers in the MRC/SSRC Unit differ in their working environment in a number of ways from their colleagues in the Psychology Department. However, this does not entail that the general pattern of their information use will necessarilly differ from that of those working in the

more conventional environment of the academic department. A comparison between the information seeking activities of the researchers in the Unit and those of the psychologists working in the Psychology Department, therefore, is of particular interest, both in terms of any similarities and any differences which may exist between the two groups.

Chapter Five

Psychology

The Department of Psychology has fifteen academic staff working on a wide variety of research topics, many of which are interdisciplinary. The Department has joint faculty status in the Faculty of Pure Science and the Faculty of Social Sciences of the University, it is associated both with the MRC/SSRC Social and Applied Psychology Unit, and with a newly established Artificial Intelligence Vision Research Unit.

Starting

The value of informal contacts when starting in a new area, which had been remarked on by researchers in the MRC/SSRC Social and Applied Psychology Unit was also stressed by one of the psychologists interviewed. He considered that this was the more effective way to ensure that up-to-date information was being obtained because of the time lag involved in journal publication.

> 'The practical way to start is to talk to people who are in the area. When I got into the stress work - the more physiological aspects of stress - this was quite new to the people in the MRC Unit here. I started looking at the literature, abstracting, picking something up, chasing references. Talked to, rang up, met people that I knew...It's not uncommon to write to the person who did the work because any work that you read up - even though it might be in the current issue of a journal is at least three years old usually' [Psychology 6A 2371]

Another reinforced the point, also made by researchers in the MRC/SSRC Unit, that it was unusual to go into an area knowing absolutely nothing about it

> 'Obviously it depends how much you knew about it. It would be rather unusual to go sort of

> straight to an area without knowing anything at all about it' [Psychology 4A 457]

These two characteristics - knowledge of informal contacts and familiarity with the field - give the established researcher or lecturer an advantage over the newcomer starting on a relatively new topic.

The research student in psychology is not so likely to have extensive contacts or long experience of working in the field, and those responsible for supervising research students saw as part of that role that the student be guided to pertinent material. The induction of research students was undertaken in a similar way to that of the MRC/SSRC Unit. The ethos or style of guidance could vary between a 'liberal' approach in which the student was encouraged to gain an overview of the different perspectives, and the 'jesuit' where induction into a particular way of approaching the topic was considered a necessary pre-requisite to any criticism of it.

> 'In the case of a research student you might give them quite an extensive reading list. The important thing there is that you must cover all the main theoretical points of view and theoretical perspectives. You must give the student an impression of a good range of the the important empirical papers, and, of course, give the student a good idea of the relevant journals and the relevant library facilities' [Psychology 1A 565]

> 'The basic idea with postgrads - there is no rule for success in monitoring postgrads - is to give them some project which they can get their hands dirty, some reading, ample discussion, and prejudice them early...Basic catholic situation - the jesuit thing - give me the child and I will give you the man. I'll say to them I don't mind you criticising my research, that's a great idea - but you are going to have to do it informedly. You had better read this stuff then you'll know what I'm talking about...This is what is going on - this is the paradigm. Induct them into the paradigm. That's reasonable - never works like that of course' [Psychology 3A 321- 360]

The value of presenting the student or novice researcher with some kind of overview or perspective on the work done, rather than a collection of 'relevant' references was also mentioned by those who had not had supervisory experience.

> 'I think I'd recommend various chapters...in books, interesting articles, the ones which I think give the most incisive summary of the most recent work. One doesn't just want a bland review, one wants something that's willing to take a more evaluative style, saying which ones are the best and what the pro's and con's for each one are' [Psychology 4A 169]

The psychologists did not specifically mention the use of library, catalogues, abstracts and indexes for starting, although one had made extensive use of two bibliographies, one which had been publis ed as a monograph by one journal, the second published as part of another journal. These had been identified by the psychologist early on in his research and, to a large

extent, saved him the effort of having to compile such bibliographies himself.

> 'There have been a number of bibliographies of statistical papers. For example, three years ago one of the British journals produced a monograph that was - as far as possible - a complete bibliography of all that had been written on applied non-parametric statistics since the year dot. It was a monograph of the Journal of the British Mathematical and Statistical Society compiled by Bernard Singer About the same time in an edition of Psychological Bulletin - a journal I use a lot - there was a bibliography published of all papers on quantitative methods that had been published in that particular journal since the thing was started. They were pretty complete bibliographies too when they were published' [Psychology 7A 026-057]

Chaining

For the psychologists, as for the researchers in the MRC/SSRC Unit, chaining was a very common information seeking activity. All the psychologists employed backward chaining from the footnotes and the bibliographies of material consulted.

> 'Generally its reading one paper and I follow it up through references at the back which generate more references, and then look out for anything that comes in' [Psychology 2A 545]

> 'Following the footnotes, the bibliography at the end, very much a snowball technique' [Psychology 9A 2077]

Forward chaining was, again, less frequently mentioned, but was employed by some of those interviewed, and considered valuable when it was used.

> 'From time to time you come across a paper that seemed to you to be interesting but after a few years you discover that there's nothing else - or you don't notice that anything else has been done on it. So you go down to Science Citation Index and see if anybody has actually cited it. I wouldn't say I used it a lot - but it's very useful when

you do use it' [Psychology 1A 642]

> 'In the course of my other reading I'd come across some stuff that people had suggested - so I would have a key article to look up in the citation index to see who's been referring to that' [Psychology 4A 503]

The lack of awareness of the concept of forward chaining was commented on by one of the psychologists who did employ it. He noted that many people had not grasped the idea of forward chaining, and that it was necessary to make explicit reference to it if the students were to consider employing it.

> 'It's still necessary, for example, to point out to students the importance of the citation indexes, the idea of working forward in time as well as backward at the end of papers is not one that people have always got hold of' [Psychology 1A 565]

Two of the psychologists interviewed reinforced the point made by several of the researchers in the MRC/SSRC Unit that citation practices involved a large element of subjectivity. Individual perceptions of what is pertinent to a particular topic can vary, and different approaches to a problem may lead to different bodies of work being thought relevant and requiring citation. It was also remarked that authors often cite material which is relatively idiosynchratic to their way of viewing a problem.

'If you take language development, there are so many ways into that that you'll find a few standard references which inevitably crop up. According to the slant of the person's research interest who's writing it, the theoretical approach that they take, and their particular inroad into language development you'll find more esoteric references on the particular aspects which are idiosynchratic to them' [Psychology 2A 375-401]

'In any paper there will be references which are relatively idiosynchratic. In the sense that they are not necessary either to the established background to which the work is done nor to the discursive interpretation into which the work is put. They may reflect some quirk in their interpretation where they drag something in which is quite different from the basic sort of thing. Thats not always the case' [Psychology 6A 2266]

Another point made by the psychologists, which had also been made by the researchers in the MRC/SSRC Unit, was that the psychologists would expect to be differentially familiar with the references in material they read. The more the topic was directly related to their field of interest, and the longer they had worked on the topic, the more likely they were to be familiar to the references to material on it. The psychologists were less likely to be familiar with references to subjects which were of more marginal interest, or which dealt with techniques specific to the citing authors work with which they were unfamiliar. Nor would they be particularly interested in following up such material, unless for some reason it seemed germane to their own

work.

'If its an area I'm working in I would expect to know and have read all the references that were really germane...For example,I would be astounded to discover a paper that was a year old which was central to the antibody work or to the anxialetic work, I would be less astounded to discover one dealing with diabetes and alcoholism. Fetal alcohol syndrome - I'd be surprised - I'd have expected to pick them up one way or another. The stress work I'm less familiar with - although as far as I'm aware we've got a cover on all the fundamental work that's been done up to the present...but I would be less surprised to discover something there because we've only been in it a year or so and one doesn't have as much contact with other people in the area' [Psychology 6A 2266-2297]

'In my previous research interests - which were on memory - then I would have been surprised if there had been anything more than a couple of years old which I hadn't seen. But the computer aided learning literature, there is so little that the references are often to rather odd things which I certainly haven't seen, and when you get into the computing literature a lot of it is internal reports - MIT internal report number 73 or whatever - which in general I haven't seen at all, and won't have heard of it because you don't get that type of thing through the Current Contents. It's much more word of mouth in the computing literature, they don't bother so much going through the established journals' [Psychology 4A 213]

'Now when I look at his reference list here I almost understand none of it because its about number studies. The next paper - I know that, and that, and that, and that. I've read that. I know that, I know that, I know that, I don't know that, I know that, thats old stuff. I don't know anything about this, this

> is a mathematics text, the next reference has got nothing to do with my thought on vision research but its a technique which this guy's using. So, on that, I've read most of the papers on stereopsis, and surface interpolation, and almost none on the mathematics which underlies this guy's method of doing surface interpolation' [Psychology 3A 107]

Bringing the activity to a close was similarly subjective. One of those interviewed pointed out that the amount of chaining required to be confident that the important material had been covered varied greatly between different subjects. Another, in the context of discussing how far he would feel he had to follow up references for a research project he was supervising, considered that he would merely need to know where the 'branches' were leading, but that the research assistant should be following them to the 'tips'.

> 'Generally its reading one paper and I follow it up through the references at the back, which generate more references and then look out for anything that comes in...It goes on for as long as it takes, because you've got to get the major people in, so it might only be a couple of authors in one area but there might be twenty in another. But I would be selective about starting at the beginning with the most recent papers, unless there is something that was a change in the viewpoint...a few years back. But on the whole I attempt some limit by starting with the most recent. [Psychology 2A 545-657]

> 'I will follow it to the point where I can talk intelligently with my research assistant about this. I would expect her to follow

> these branches right to the tips - I would know where the branches were roughly leading' [Psychology 9A 2105]

Several of the researchers at the MRC/SSRC Unit had mentioned that repeatedly seeing the same references cited provided some confidence that the search had been reasonably comprehensive. One psychologist also considered that repeated corroboration, or changes in the level or type of references cited also provided grounds for thinking that the search could be safely brought to a close.

> 'Where you get two separate references coming up which begin to corroborate the story, and it looks as if the third and the fourth would do that. Or that another branch you've gone into is getting more detailed - you're getting into the psychopharmacology of brain action which is not our concern. So corroboration and the boundaries of relevance I suppose' [Psychology 9A 2116]

Browsing

The psychologists, like the researchers in the MRC/SSRC SAPU, made relatively little mention of browsing, and, they, also, made heavy use of Current Contents. Possibly, use of Current Contents, because it greatly eases the task of scanning a large *number of* journal contents pages, serves to integrate aspects of browsing and monitoring, or, at least, make it difficult to clearly distinguish between the two. The one activity - scanning Current Contents - being employed both to monitor the contents of some journals, and to browse or quickly scan the contents of others.

> 'Journal of Education and Social Work - I would have a low expectation of there being much in that. British Journal of Criminology, probably, there's likely to be something in there on environmental psychology. Studies in Family Planning - er no! One has expectations of particular journals - something like the Journal of Polynesian Society one would look at for sheer fun...Sociological Review I would scan simply because it's Sociological Review - it's standard. All of the psychological things I would scan' [Psychology 9A 572-595]

Differentiating

Differentiating between sources according to their substantive topic was as much a feature of the information seeking patterns of the psychologists interviewed as it was of the researchers at the MRC/SSRC Unit. The psychologists frequently had a very good perception of the sources which they would expect to publish material in their areas of interest.

> 'They're the ones that publish papers on development first of all. It could be child development or it could be development more generally. There are certain journals that tend to specialise in that...The more general journals like the British Journal of Psychology don't actually publish much in the developmental way. So although I'd probably just check - skim over the contents page of the British Journal of Psychology when it comes out, every now and again, I don't normally expect to find very much' [Psychology 8A 091-094]

> 'All the specifically language journals which I suppose there must be - taking child language and linguistics - about fifteen. There are the educational journals, of which there aren't many, say Research in Teaching English...There's American journals - which I suppose there are about four. There are then all the more general journals - like Memory and Cognition - or say things like Child Development which cover developmental issues.

> I should suppose there are about thirty journals altogether, of about which I suppose ten I would consider very important, and have to be looked at in every issue, but general journals like British Journal of Psychology are quite likely to contain something as well' [Psychology 2A 131-147]

Also mentioned was the influence which a change in the editorial policy of a journal could have on the probability of that journal containing material which was likely to be pertinent.

> '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, and a lot of things of this sort...Some of them are, shall we say, slightly more hard-nosed than others - but there's a possibility that you'll get something in all of them' [Psychology 4A 052-065]

Differentiating by approach or perspective was particularly important for one of the psychologists interviewed as the basis of his research was the applications of statistical techniques to problems in psychology. In his case the approach was more significant than the substantive topic, and he differentiated sources primarily according to the likelihood of their containing material concerned with appropriate quantitative methods.

'On the psychological side Psychological Bulletin and on the statistical side, Biometrika, Psychometrika, Biometrics, and so on...Sometimes the Journal of Mathematical Psychology. It's perhaps not fair to mention that it tends not to be very suitable...Because it tends to be less on statistics and more on mathematical models of psychological processes, mathematical theories of psychology...Psychological Bulletin is the sort of journal because that's one that has articles on quantitative methods, and is read by psychologists in order to increase their awareness of quantitative methods' [Psychology 7A 097-123]

The significance of differences of quality between sources was stressed by one psychologist who, as part of a research project had had to examine a very large number of journal sources. As a result, there were certain journals which he would, as a matter of principle, not follow up material in.

'I don't think some journals perhaps review their articles very well. I really only found this when I out going into journals which I didn't normally read, when I started looking for articles on written language. So there are some journals now which I never bother to send for articles for from Current Contents - I don't know if you can put an exclusion of certain journals in a search' [Psychology 2A 2145-2154]

As with some of the researchers in the MRC/SSRC Unit, not all the psychologists found it possible to differentiate a limited number of sources in which they would expect material to appear. One psychologist made this point with reference to the medical literature.

'Medics publish in the most extraordinary things. I mean you will get something which is about, let's say, noted Vietnamese use of drugs compared with American soldiers use of drugs, published in the New England Journal of Medicine. You would never have thought of looking there - but I think that was a particularly broadly scattered thing. I'm also aware that a search like that doesn't scan or even have any way into say the Journal of Asian Studies or the Australian Journal of Criminology or something like that' [Psychology 9A 689]

Monitoring

A major way in which the psychologists interviewed monitored their fields was through the journal literature. This was done, in part, through monitoring the actual journals themselves, but also through the use of Current Contents. Several of the psychologists interviewed were particularly heavy users of the Current Contents service, and relied on it as a

principal way in which they monitored the literature.

> My intention is to read journals as they come - because I take quite a number of them...There are periods, of course, when it's quite impossible, because there are so many other things to do, and they tend to then get piled up' [Psychology 2A 203-216]

> 'One's immersed in it. You go through things like - routinely Current Contents, routinely glancing through journals as they come across your desk. I don't do it systematically' [Psychology 6A 580]

> 'I go through Current Contents all the time...I also survey Science, and Nature, every now and again articles come up of relevance there' [Psychology 5A 315 480]

In their use of Current Contents the psychologists differentiated between the Social and Behavioural Sciences service and the Biological and Life Sciences service. Some used both equally, others would use both but expect to find most of the pertinent material in one or the other. Which of the two was preferred depended on the nature of the individual's research interests.

> 'The stuff on most things, the stuff on memory, there are various journals I look through periodically. Every two or three months I go down to the library and have a look at the recent issues...The main source of information is Current Contents, which we

get, the social sciences Current Contents and...life sciences, the life sciences has some stuff but its not usually relevant, the social sciences is much more relevant to me...My major thing is just writing off for reprints from Current Contents...Current Contents arrives on my desk...There's a section on psychology and so I look through the things in that. There's also a section on education, and I'll look through that, but with less expectation than the psychology one' [Psychology 4A 026-046 075 078]

'Looking at Current Contents. I often look at the statistical journals...Reading certain journals. On the psychological side Psychological Bulletin, and on the statistical side Biometrika, Psychometrika, Biometrics, and so on, just general reading, and so on, keeping up to date with what was coming through...I haven't used Psychological Abstracts for ages. I have my work cut out keeping up with what comes through in Current Contents without having anything further...They tend to come up in the social sciences issue and the biological and life sciences issue. The general journals such as Psychological Bulletin are covered in both. Biometrics, Biometrika, Psychometrika I assume just come in life sciences' [Psychology 7A 068 097 433 470]

'All I've done so far is to keep an eye open for articles...in Current Contents. I think that the journals that I routinely cover in the development area will have picked up 90% of them anyway - Child Development, Merrill Palmer Quarterly, Developmental Psychology, and so on...The social and behavioural, and the life sciences. I use both. The social and behavioural one is more relevant. But the life sciences one has Nature and Science in, and a few animal behaviour type journals in, which aren't in the social and behavioural one' [Psychology 8A 098 296-297]

'I do things via Current Contents. Fairly systematically through Current Contents life sciences and Current Contents social sciences. Social sciences is, hands down, the best, except for the drug area which is better covered in life sciences. (The social sciences) is much broader canvas of things...I will scan in that sociology, anthropology, social issues in philosophy, psychology, begin to tail off in psychiatry, rehabilitation and special education picks up the blind, education possibly, geography planning and development certainly...So Current Contents is a major source' [Psychology 9A 440 452]

'What I do mainly is to read Current Contents pretty regularly. Both social sciences and the biological life sciences. So I can go through that and thats a pretty comprehensive coverage. Plus when in the library just going through a run of a relevant journal such as Cognition, or British Journal of Developmental Psychology, or whatever. Plus going to Index Medicus or Psychological Abstracts when one wants to to sort of follow up some particular point thats arisen' [Psychology 1A 206]

Informal contacts were also mentioned by the psychologists as representing a valuable way of keeping up-to-date with developments in their areas.

'You tend to rely also on a certain amount of informal contact. Knowing people who are in the same field saying have you read so and so, have you not read so and so, and have you seen what so and so has come out with. We keep each other up-to-date to some extent' [Psychology 1A 689]

'By and large you keep each other alert.

> Mainly about stuff which is - you find out of the way in journals you wouldn't normally be looking at' [Psychology 6A 2320]

Where a person has been working in the same field a considerable time his present work may be based on a long history previous work carried out in that area by that same person.

> 'You're talking about work based on work based on work based on work that goes back twenty years. I did my first experiment with librium in 1962. So that particular sort of work I have been doing off and on over the last twenty four years. The antibrain antibody work has been going on over the last six to seven years - but again the behavioural side of the measurement we have been doing for the last twelve to fifteen years. I started working with behavioural genetics eighteen years ago - 1965' [Psychology 6A 375]

Extracting

Extracting from journals was mentioned by several of those interviewed, as was the use of abstracting services as a substitute for chasing through the individual journal series.

> 'Actually pursuing particular journals through' [Psychology 9A 2077]

> 'Abstracting services - you can go forward in the literature they're sampling. So that basically it saves you the problem of going back in individual literature runs' [Psychology 6A 2511-2520]

In a related way, where monitoring had lapsed, due to pressure of other work, a retrospective search through the material accumulated could be employed to bring the person up-to-date.

> 'My intention is to read journals as they come - because I take quite a number of them - and to read reprints which are a very valuable source of information. There are periods, of course, when it's impossible because there are so many other things to do and they tend to then get piled up. When I'm desperate, and I haven't got a lot of time I flick through the index and skim through the

articles that I know are important and then go back to them later. [Psychology 2A 203-218]

Conclusion

Despite the similarity in broad disciplinary base between the researchers in the MRC/SSRC Unit and the psychologists working in the Psychology Department the working environment of the psychologists was very different from that of the researchers in the MRC/SSRC Unit. The psychologists were working only part-time on research and so had to accommodate their research projects with the demands of teaching, and some had to cover a broader span of subjects for their teaching than for their research.

There was some similarity of interests and research collaboration between some of the

psychologists, and some also were responsible for research students who might be conceived of as representing a small research team, but, in general, there was a greater diversity of research interests between the psychologists in the Department than between the researchers in the Unit, and more individuals were working separately on their own research interests. In addition, the induction of research students or assistants was more sporadic. The diversity of interests meant that there was less scope, internally, for cross-fertilisation of ideas and information, and, of course, the Head of Department could not operate in the same integrating and coordinating role in relation to the Department's research activities as the Director of the MRC/SSRC Unit could.

Nevertheless, the information seeking activities of the two groups are similar both in form and content, and both can be accommodated within the overall framework of the behavioural model.

Chapter Six

Education and Continuing Education

The Division of Education and the Division of Continuing Education, together with the Department of Information Studies, form the Faculty of Educational Studies. The main teaching activities of the Division of Education centre on the Postgraduate Certificate of Education and other taught Diploma, higher degree, and in service courses. The research interests of the twenty staff cover most aspects of educational theory and practice. The Division of Continuing Education

offers a diverse range of courses for adults and the research and teaching interests of the sixteen staff reflects that diversity of interests.

Starting

The experience of the educationalists was similar to that of the researchers in the MRC/SSRC Unit and to that of the psychologists. Several mentioned that if possible they would start with informal contacts or from working from what they already knew, in either case the intention being to identify a few key references which could serve as starter references.

> 'I think one of my first steps is normally either phone up, or write to, or go and see a colleague I know who has done more work in this area than I have and ask their advice on some of the key books to start looking at, and gradually broaden out from there' [Continuing Education 10D 360-367]

'I would be most unlikely to to start by going through British National Bibliography or approaching in that way. I would probably rely on having noticed some books in the area already, that perhaps have been published in the last two or three years, but I can't imagine myself going into a completely new, unknown area of teaching. What I would do is is I would begin by examining what I already knew, and what my present perspective on it is' [Education 4D 366-384]

In a similar way, many considered that they would direct new researchers in their areas to key books, journals, or journal articles, which could serve both to introduce the person to the area and provide references to follow up. One of those interviewed also mentioned that if there was an up-to-date bibliography available he would also direct a newcomer to that.

'There are certain keynote books that I would want to refer them to. I would refer them to the British Journal of Guidance and Counselling and some keynote articles in it' [Education 6D 573]

'I'd hand them a book, an up-to-date book, on the subject...a kind of textbook type of book - that reviewed - that was fairly up to date...I would hope that I would know at least some book that was up to date, so that I'd ask them to go and have a look at the book,and following that I'd ask them to go and look at journals - look at the recent editions of journals in the particular area' [Education 7C 481]

In contrast to the psychologists, and to the researchers in the MRC/SSRC Unit the educationalists

did not mention making much use of reviews or review type material when starting, and only one mentioned using the library catalogue at this stage.

> 'Just going down to the library and using the library catalogue myself, and using a subject index and author indexes' [Continuing Education 9D 236-251]

However, several had had computer based searches undertaken by the library. The comparatively large number of educationalists who had had such searches undertaken was probably due to the interest of the education librarian in providing, and promoting, such searches, rather than reflecting a *particularly strong* interest in computer based searching on the part of the educationalists.

In general, those who had had such searches carried out had requested them when they were starting on the topics they wished to have searched. Often, the search was in an area with which they felt unfamiliar, or certainly less familiar, than with their other research interests. Sometimes they had a prior conception that the material was likely to be widely dispersed in the literature, and also to be in a literature of which they might not normally be aware. The motivating factor for having the searches carried out seemed to be that they were perceived as representing a quick means of starting, rather than any belief that the computer search constituted a

sophisticated alternative to their other ways of searching. The great perceived advantage being the speed with which they could obtain a basic set of references from which they could then work.

> 'That literature search was to get me going in a field which I hadn't specifically looked at, and which was, I was thinking might in fact draw upon journals which I wasn't familiar with in the education field. Which, of course, it did. Which I normally didn't look at because...it was straddling the two disciplines. I know the politics stuff, but the more specifically educational stuff I wasn't as familiar with' [Continuing Education 10D 170-183]

> 'The one on machine intelligence was an area that was, is, beginning to become quite important in the philosophy of mind and I wanted to get into it quickly' [Education 2D 288-300]

> 'Its a newer field to me than scripted drama...I was feeling my way more, and I wasn't sure in fact what had been written on the subject. I knew that very little was published in England and I wanted to see what the Americans had done' [Education 1D 112]

In their comments on the individual searches the educationalists highlighted particular difficulties with terminology, with differentiating the level and type of treatment, and with the amount of marginal or irrelevant material brought up by the searches. Overall, the perception of computer based searching was

that it was a relatively unsophisticated means of searching for material, but could be useful at an early stage when the searcher had little to go on.

> 'With the best will in the world you cannot classify materials under simple descriptors in a way which will allow a person looking for material to separate the gold from the dross. You tend to get a pretty long printout which you then have to plough through to what is actually relevant and what is not. I don't see any way round that - its not a criticism of the system its just the way the world is' [Education 5D 290]

> 'You would identify a large number of false positives. What I mean is that you would incorrectly identify some sources as useful...that may not be, turn out to be, all that useful, because they may not answer questions, they may not relate to my constructs of a field, the way in which I might make sense of a field of education. Take parental involvement...You can come across titles, particularly in the american literature of books on involving parents, which, just from the title and the publisher you have no way of knowing what the book is intended for. Whether its intended as a research report or a guide for teachers or what, so you can have that kind of problem' [Education 4D 407-454]

> 'Obviously, if you are totally stuck you will use a peripheral thing, again if it's easily accessible, but an awful lot of the stuff, again, was in the the States, in fairly obscure journals in the States, by the time you've tracked them all down and chased them all up it's next year' [Education 8D 296]

'I think that pretty typical of most online searches - that you get an awful lot of stuff that is either irrelevant or marginal - but it's useful to know...The point of an online search is that it cuts down the labour of going through all the journals' [Education 6D 500]

Chaining

Backward chaining was as common a feature of the information seeking activities of the educationalists as it was of the psychologists and of the researchers at the MRC/SSRC Unit. The backward chaining could take place either from footnotes or the bibliography of the citing work.

'Footnotes, and so on definitely' [Continuing Education 10D 144]

> 'Obtain references from the bibliography' [Education 4D 144]

> 'I suppose you pick up a lot of references originally in other peoples publications in their footnotes, and so on' [Continuing Education 11C 041]

One of those interviewed drew particular attention to the importance of the context of a reference in the decision whether or not to follow it up.

> 'When you come across something to which reference is made in an article, the reference is made in a context with, if you like, an annotation, which is very much more informative than a title and a source. After all the vast majority of things to which reference may be made in an article you conclude that you don't want to see for the same reason' [Education 5D 326]

However, in contrast to the psychologists and the researchers in the MRC/SSRC Unit, none of the educationalists made any mention of attempting forward chaining, and most had no knowledge either of the existence of the Institute for Scientific Information's Citation Indexes or of how they could be used.

Browsing

In the context of browsing, there was also a contrast between the information seeking activities of the educationalists and those of the psychologists and researchers in the MRC/SSRC Unit, in that the educationalists did not make any use of Current Contents. The reason for the lack of use of Current Contents was not clear: some made reference to a similar type of service offered by the education librarian, who circulated photocopies of the contents pages of the journals received in the education section of the library, and several mentioned browsing through the journals themselves in the library. In this respect, browsing through current journals may also be seen as making an important contribution to maintaining current awareness, representing a diffuse form of monitoring.

> 'I occasionally go and browse through the journals on display in the main library - in the Education section. There are about fourteen or fifteen on display and I might look at four or five to see what was in them. I wouldn't do it very systematically [Education 3C 150-173]

'Just by going and having a look sometimes. You can use lists of journals that you have in various places - like on the ends of the stacks. Basically you'd go down there, find something that sounded relevant and you'd skim down the indexes for the year...when you've found one thats interesting or fairly specific you'd either sit down and read it or else get a photocopy of it' [Education 8D 034-029]

'I just went down to the library and looked around and just saw it (The Journal of Children's Education) on the shelves as a matter of fact. I knew there must be such a journal though, in fact I probably did sub-consciously know about it, because I read the Times Education Supplement regularly, every week, which is the way I keep up to date with whats in education generally, and I probably somewhere along the line saw a reference to it - I knew that such a journal existed anyway' [Education 7C 200-215]

'I often scan journals as they arrive in the Education Library. Once or twice a term I'll stop and look at current journals, and if anything appears interesting I would follow it up' [Education 4D 138]

The activity of browsing is not confined to browsing collections of books and journals on library shelves, wherever there is some collocation of like material there is the possibility of browsing. One educationalist reported browsing through an abstracting service, but did not consider it a particularly significant feature of his information gathering activities. Another found that a valuable feature of

conferences were the publishers' displays of books which could be browsed through for interesting material.

> 'Sometimes I will leaf through the pages of Sociology of Education Abstracts. I occasionally mark some - I tend not to - I don't sit down and think that I ought to put that in my card file regularly enough to convince myself that it's important' [Education 3C 562]

> 'One of the things...that I do get out of conferences is that often they actually do have bookstalls there, whereby they get all the publishers to put their newest offerings up. I find that very helpful, partly because I can actually pick up the books and see if they can be useful for teaching purposes' [Education 8D 415-419]

Differentiating

Differentiating was employed by many of the educationalists. Several of those interviewed mentioned employing distinctions or divisions in the journal literature as a means of differentiating sources of interest. This could be in terms of the substantive

topic covered by particular journals.

'Mainly in the journal - Science Education, and the European Journal of Science Education. I keep dropping into the library to see if the next one has come in. My other main regular source is the School Science Review - which I get myself...When you're interested in science education there are not that many other relevant journals. I do follow up other journals, I get the Journal of Curriculum Studies for example, and I have a quick waffle through the contents pages of probably 15 others whenever I go into the the library. But it is these journals that I would definitely take off the rack before I even look at the contents columns. Because a fair proportion of what is written in them is quite central to the issues that I'm interested in' [Education 5D 159-170 199-212]

'Political Studies, Policy and Politics, Local Government Studies, these are the main ones, but there are also Public Administration and one or two of the sociology journals as well, and sometimes Town Planning and other journals of that kind...They are the journals in the field - the political science journals' [Continuing Education 10D 030-043]

'There's only one specifically devoted to pastoral care and that's a relatively new one - it's called Pastoral Care in Education. Then there's the British Journal of Guidance and Counselling, Educational Research. There are one or two American journals...There are only really a few that are concerned with this area. More recently there have been odd articles in the British Journal of Educational Research, one might find the odd article in Educational Review, and in Research in Education, but they are pretty scarce' [Education 6D 100-141]

Differentiating by approach or perspective was also mentioned as an important criterion for discriminating between different sources. One of those interviewed mentioned differentiating in this way with respect to the broad disciplinary base of the journals. Another differentiated between sources in terms of the particular type of approach or perspective on the subject adopted.

> 'Nearly all the philosophy journals. There are two or three philosophy of education journals...and unfortunately a number of the generalist educational journals carry articles in philosophy...Just straight philosophy, I would say there must be about forty. I wouldn't look at all of those...There would be eight to ten, at least, that would be central. I would expect to find at least one article in each month or quarter. For example, the Journal of Philosophy nearly always carries something of interest' [Education 2D 124-151]

> 'British Journal of of Sociology of Education, British Journal of Educational Psychology, Educational Research, Journal of Education and Teaching, Forum, Scottish Educational Review, British Journal of Educational Studies. They are the ones that might publish things that I see as part of the questions that I'm interested in. The others are less likely to, they're a different paradigm, they're not interested in exploring questions in the way that I'm interested in exploring them' [Education 3C 146-173]

Several of those interviewed made a distinction between 'academic' and 'practitioner' journals, on the

basis that they would expect different types of material to appear in the different types of journals. This is probably a key distinction to make in any field which has a practitioner as well as an academic dimension. The 'practitioner' journals might be further sub-divided according to the different professional groups or associations involved. One mentioned working from the level of news or topical coverage of subjects through to more theoretical approaches, this was related to his particular interest in examining ideas, within his area of research, which had entered into topical debate.

> 'I do make a distinction between what I call practitioner journals and academic journals - but in my field in education there are a lot of journals that people read when they come on course which are practitioner journals - they are the journals of remedial teachers, and the journals of teachers in special education - and they are definitely distinct, in many ways, from journals like the the British Journal of Sociology of Education, or the British Journal of Sociology, or Sociological Review...The articles have a more theoretical orientation in the the academic journals rather than the practitioner journals which are concerned with pragmatic issues' [Education 7C 560-576]

> 'I would use Adult Education - although that tends to be a bit academic - then there are some of the occupational psychology journals, then there are two or three journals fairly specific to the area of educational psychology, which are the journals of the professional associations I suppose. They have the occasional article, not all that many. Then there are the journals connected

with the further education, which tend to be the journal of what was the National Association of Teachers in Further and Higher Education, again they tend to have odd articles in the areas relevant' [Education 8D 057-071 535-543]

'On the industrial relations side of things...what I need to use relates a great deal to what is topical, and so I'm likely to work as much from the press and publications like Marxism Today - and then work back to the academic literature' [Continuing Education 9D 037-072]

Monitoring

Monitoring the pertinent journal literature was a major way in which many of the educationalists kept up-to-date with their field. Usually they had identified a limited number of journals to monitor by prior differentiation.

'I keep up-to-date with journals like the Agricultural History Review, the Economic History Review, the publications of the various archaeological societies, and so on...I keep very much up to date in the journals like the Agricultural History Review, that's the one of greatest interest to me...Economic History Review I only read the articles that are of relevance to me, whereas in the Agricultural History Review I'll read it from back to front because I'm interested in what they are doing and this sort of thing' [Continuing Education 11C 186-267]

'Mainly the journal Science Education, and the European Journal of Science Education...My other main regular source is the School Science Review, which I get myself' [Education 5D 159-170]

'With labour history I've no problem at all because I edit the journal in the field which is the Bulletin of the Society for the Study of Labour History...and, of course, we review books in the journal so I've got a very good picture of everything that's coming out, even if I haven't read it I know what's coming along. Apart from that one reads a number of other journals anyway where material comes up' [Continuing Education 9D 037-072]

'I read the Times Education Supplement regularly every week, which is the way I keep up to date with what's in education generally...A couple of sociological journals, Sociological Review, the British Journal of Sociology...as well as the practitioner journals like the Association of Educational Psychologists Journal, and then various journals...tangential but cognate I suppose, like Special Education, Remedial Education' [Education 7C 200 222-233]

'There is a journal called the Industrial Tutor which I read regularly, and that's a source of information. There is another journal, an American journal, called Labour Studies Journal which I look at regularly. There's a journal called the Trade Union Studies Journal which is published by the Workers Educational Association...Then of course there are other things you have to look at, the Trades Union Congress annual reports to pick out statements of policy from the Trades Union Congress education department. Trade union journals can be of some use occasionally - though not very often - the journals of individual unions' [Continuing Education 9D 121]

'On Fridays spend the whole day in the stacks in the journals. I should like to think that I would be in the stacks going through the journals four times a month' [Education 2D 114-122]

One, whose research interests centred on key topical ideas influencing thought and policy in the field, made similar use of the press.

'I read the papers I see what people are discussing...I think that in terms of what I'm interested in publicity is quite important, in the newspapers, the Times Education Supplement, the Guardian Education page' [Education 3C 146-173 359-360]

Interestingly, none of those interviewed made any use of Current Contents, though a similar type of service provided by the library was mentioned.

'The education librarian...is now simply photocopying all of the index pages of

> virtually all of the journals relevant and putting those on successive sheets...He is simply presenting you with the titles journal by journal and you can see from the titles if its going to be worth actually going to that particular journal' [Education 8D 092-123]

Two remarked that monitoring publishers' lists represented an important way in which they kept up-to-date with recently published material.

> 'Obviously academic journals, and also on the list for the distribution of of publishers' material, the general clearing house where you can tick a number of boxes. I tick appropriate boxes and get a fairly steady stream of material from the publishers...I also read the Times regularly every day and pick up an enormous amount of what's happening currently, particularly from government sources, from there, so that I can get the government publications as they come out' [Continuing Education 10D 022-038]

> 'Publishers' lists - word of mouth' [Education 6D 215]

One mentioned regularly examining the new accessions counter of the library, another commented that he relied to a large extent on knowing of the existence of bibliographies produced by a colleague working at Cardiff. He knew that, if necessary, he could use these to bring himself up-to-date on a subject within his area of interests.

> 'On Monday morning - as I don't teach that

morning - I go to the library and check the new accessions, the book accessions on the desk, and take what I need from there' [Education 2D 114-122]

'I'm very aware of a colleague who works at the University of Cardiff - University College Cardiff - who produces very lengthy reviews of the state of the art...and usually these have very long bibliographies at the end which in a sense can serve as updating for me if I want to update myself' [Education 3C 188-197]

Two remarked on the value of abstracts of conference proceedings, select bibliographies, and book reviews published in journals.

'The British Psychological Society Bulletin generally has abstracts of conference papers...It has abstracts of papers delivered to annual conferences and regional small conferences. I normally read this - I mean I normally look through them and occasionally there would be something I'd actually make a note of, perhaps follow it up, write to the person concerned, or make a note to look out in journals for what they are doing' [Education 4D 330-338]

'The Agricultural History Review and the Economic History Review also publish and annual list of articles that have been written that are relevant to the subject, and you find that some of the regional journals do this as well. They also contain book reviews which are helpful' [Continuing Education 11C 186-267]

Extracting

Extracting from key journals was mentioned by several of those interviewed.

> 'I thought I'd better look at management education, so I had to spend quite a lot of time up at the Crookesmoor Library going through journals that had to do with management education' [Continuing Education 9D 181]

> 'Basically I'd get hold of the copies of the past five years of the major journals. Just have to read through the articles which are relevant...Then from there follow up for specific purposes from the the actual references that you get' [Education 8D 027-039]

> 'There are one or two British Journals that I would probably look at for more general work on job satisfaction. The British Journal of Occupational Satisfaction for example, maybe Personnel Management...I would look - go through - the index and the contents pages first of all and identify articles that seemed to be promising from the titles' [Education 6D 500-518]

Monitoring and extracting were closely related in the activities of two of those interviewed. One had not

been monitoring the journals in an area in which had once written, and was now intending to write again. He considered the first step would be to go back to the relevant journals and carry out a retrospective exercise in extracting from them. The other used extracting as a primary means of keeping up-to-date, by each year systematically working through the last issues of all the relevant journals in his area.

> 'I did write a piece in 1978 for the American journal that I mentioned Labor Studies Journal and I'm probably going to take that as my jumping-off point - where I finished with that piece. But between now and 1978 since I've not really had to write anything about it I've not been watching all these journals and so on very carefully...And when I sit down to write it I'll spend some time first of all preparing the ground by going through all those journals' [Continuing Education 9D 162-181]

> 'Now I've got a general updating system that in the Summer of each year - when the teaching load gets a little bit quieter - I'll typically go into the library and scan the last year of all the journals all at once...I'd actually physically get the journals. Go down the actual indexes and check out particular things, and when I came to titles which appeared relevant for any of these I would read it then and there or I would make a note of it or photocopy it then and there - depending on how much time I'd got' [Education 8D 135-149]

At other times extracting was less planned and might proceed from browsing, then using indexes to see if it were worth actually going to a particular issue

of a journal.

> 'You can use lists of journals that you do have in various places like on the ends of the stacks. Basically you'd go down there, find something that sounded relevant and you'd skim along the shelves and pick out the actual journals, you'd skim down the indexes for the year...If there were two or three titles in a particular journal that would give you more of an incentive to go and find that particular journal than if there were just one which may be relevant or may not be relevant - depending on the specific contents of it, which, of course, you can't tell from the title anyway' [Education 8D 027-039 123-125]

A feature of the extracting activities of the educationalists which was not mentioned by the psychologists and by only one of the researchers in the MRC/SSRC Unit was that of going through publishers' lists to identify recently published books.

> 'I get publishers' catalogues -I don't read them all actually, I don't flick through them all -I flick through about four or five that I know from past reputation are worth flicking through. They tend to be English ones rather than American ones' [Education 7C 259-264]

> 'You look at the publishers' lists and read your way through those - it takes quite a long time...A lot of them just get sent here. I get the drama section of the Methuen output for instance. Hutchinson's do a drama catalogue...Cambridge who published my first book send me their catalogue, MacMillan send me their's now' [Education 1D 041 046]

'I have, in the past, in Manchester, maintained, created and now maintained, contact with all the publishers that I know of in this country, so that they send me publishing material that they normally send the libraries and bookshops, the advertising new books and so on. They also send me manuscripts to comment on in my areas which is useful [Education 2D 108-112]

Conclusion

The educationalists represent a far more disparate group than either the researchers in the MRC/SSRC Unit or the psychologists. Their original disciplinary base was varied and there is a further difference in interests between those working in the Division of Continuing Education, who represent subject specialists dealing with a particular type of student, and those in the Division of Education, who, although possessing different types of subject expertise, were primarily

concerned with the professional and intellectual problems of the field of education.

The educationalists were similar to the psychologists and different from the researchers in the MRC/SSRC Unit in that they could only engage part-time on research, and, again, their teaching interests might have to cover a broader span than their research interests. The differences of research interests were more pronounced than for the psychologists and offered even less scope for collaborative research within the Divisions. Overall, the educationalists research interests were specific to the individuals concerned rather than representing the kind of team project research characteristic of the researchers in the MRC/SSRC Unit and some of the psychologists, and the induction of research students was sporadic.

There were also some significant differences between the information seeking activities of the educationalists and those of the other two groups. There was little mention of reviews as a means of starting off in an area. None of the educationalists employed forward chaining, and none used Current Contents, although, some did mention making use of a similar current awareness service provided by the education librarian. However, the educationalists did make more mention of browsing than did the researchers in the MRC/SSRC Unit and the psychologists, and many reported making heavy use of publishers' lists, a

source which was hardly mentioned by the other two groups. Some of those in the Division of Education also stressed the importance of the distinction between the academic and practitioner journals in the field.

These differences may derive from a number of sources. The non-use of forward chaining may be due to lack of awareness or lack of familiarity with the Institute for Scientific Information's citation indexes, but why the educationalists should be so much less aware of their existence, or familiar with their use, is not clear. This is also the case for the non-use of Current Contents, although this may also relate to the existence of the similar current awareness service provided by the education librarian. The relatively close proximity of the library to the Division of Education may also encourage browsing of the library collection itself.

The heavy use of publishers' list by some of the educationalists is evidence for the existence of a significant difference in the information seeking patterns of some of the educationalists, compared to those of the researchers in the MRC/SSRC Unit and of the psychologists. This would seem to indicate a relatively greater emphasis placed on book as opposed to journal materials for those concerned. However, this difference of content in the information seeking patterns, can be accommodated without alteration in form of the behavioural model, with monitoring and

extracting from publishers' lists substituting for monitoring and extracting from journals.

Chapter Seven

Economics, Economic and Social History, Geography, Politics, Sociology, Prehistory and Archaeology

The Division of Economic Studies, Departments of Economic and Social History, Geography, Political Theory and Institutions, and Sociological Studies, together with the Department of Psychology and the Department of Japanese Studies form the Faculty of Social Sciences. The Department of Prehistory and Archaeology is in the Faculty of Arts. The teaching and research activities of the members of the various

departments cover the whole range of subjects in the social sciences.

Starting

The value of personal contacts when starting in a new area, both as a source of references and advice was remarked on by several of those interviewed. Two considered that their first step would be to contact people who knew something about the topic, a third thought that he would try and read something which gave him some idea of the field or topic but would almost immediately proceed to try to contact people working in the area.

> 'I'd go to my nearest colleague in economics who is interested in that thing and ask him what are the salient half-dozen things one would start with. That's the way I'd do that catching up exercise' [Economic and Social History 4C 566]

'My first thing would probably be very much a personal thing. I'd get hold of somebody I knew who knew something about it - even now that would probably be my first step. It's one thing to have a list of references - its another one to know which are the most important. I think I'd almost certainly do that I'd get hold of a friend or colleague and get them to give me some guidance' [Sociology 10C 515-520]

'I would probably start by trying to find some very brief but obviously central references about it...and I would spend not very long on that, I think, just acquaint myself generally with the issues before trying to think of people that I know who were teaching or researching in that area, and then having stopped myself being an absolute idiot in it, as it were, I would ring them up I think' [Sociology 13C 2048-2070]

The supervisor may fulfil this role for those research students who have little knowledge of the field or who have no personal contacts in it. Both those who had had experience of research supervision and those without such experience considered that a first step in the induction of research students would be to direct them to key articles, books or sources.

'One of my former colleagues...is now wanting to do a part-time M.Phil in the area...It's a narrowly focussed thing. I wouldn't be concerned to say go and do a general reading of sociology first, I would be concerned that he knew about the important bits of work that have been done in the area, and that would not involve such a prodigious amount of reading...There is a wadge of journal articles. Or itsy bitsy things - things in Community Care are not journal

articles...There are longer journal articles too. A limited number of books' [Sociology 10C 207-234]

'I could quite easily produce 2 or 3 core articles, or the core articles on each of a number of topics such as perception, attitudes etc. That wouldn't be difficult' [Economics 1C 2116]

'I would try and direct them to more general and more specific secondary literature - published stuff - books and articles. You would also at the same time advise them to use the bibliographical apparatus in those books and articles. If you happen to know that in the field there is a bibliographical guide then you would certainly send them to that - a list of references published in a certain area' [Economic and Social History 5C 184-204]

The other social scientists interviewed did not mention reviews and review type material as being particularly significant when starting, although one did mention employing a review type journal.

'I started with the last five years Economic History Reviews, which are standard, well, you know, that is the economic historians journal - Economic History Review. I looked through that for the last five years...One looks at those things first and then there are standard specialist things like The Newcomen Society for Engineering History, and The Agricultural History Review for agrarian history, and so on. You go to those - the obvious ones to start with' [Economic and Social History 4C 462-473]

in contrast many of them had employed, or considered that their first step would be to employ library catalogues, bibliographies, abstracts and indexes when starting on a new topic.

> 'I'd probably go down to the library to use the catalogue - subject index. I'd ask people in the Department...I'd probably try and find something from the catalogue index on the shelf and then use the references contained therein to try and follow it up' [Economic and Social History 5C 487-491]

> 'There are some literature surveys in the field, including collections of abstracts, and I used those. There is a good abstracting service - a thing called Social Services Abstracts - and I used that...I think I did start from the abstracting area really, also I suppose I started from what I knew to be the key texts in the general field of child care' [Sociology 13C 091-128]

> 'When I started upon the cave climates I spent quite a lot of time working through previous issues of the Geophysical and Astronomical abstracts which the American Metereological Society produce monthly...I've not really looked at the abstracts since probably 1980-1981 sort of time (November 1983). When I started I had very little background...and I must admit I hadn't much experience at that stage in knowing what it was like. So I had to start what I felt was from scratch' [Geography 7C 340-407]

> 'I've used indexes a bit more often - as distinct from abstracts - things like British Humanities Index. That would again probably

> be background reading. There might be two things - first of all would be a trawl for ideas, and the second thing would be as a way into a new area. Obviously there's an overlap between these two but they are a little bit different' [Sociology 11D 250-257]

One of the sociologists - who had had a computer based search carried out by the library - expressed a similar opinion to those of the psychologists and educationalists who had had such searches carried out. The computer based search was not perceived as a particularly sophisticated alternative to other means of information searching, and significant problems were experienced with terminology.

> I'm much happier and feel that I can do a better job myself plunging into the literature and simply tracing it myself through the bibliographies of other works. I certainly would not trust a straightforward literature search produced by these computerised bibliographies. I'm not blaming them for that - I mean it clearly does relate a lot to the nature of the topic you're looking at and your ability to describe it or to identify keywords. Now, when you're dealing with the keywords 'elderly', 'dependency' and let's say 'institution', then you're tapping a very big literature, so its not surprising that you end up with a lot of stuff...but but the thing that always amazes me is how little comes out that is central to your own interests' [Sociology 11D 161-193]

Nevertheless, such searches were perceived of as having value, in that, even if not particularly successful or complete in themselves, they could be used to very quickly identify a number of starter references from

which someone could begin chaining.

> 'I wouldn't write off the kind of service that the library does offer...What I'd use it for was to identify those - what might be quite few books or articles - that would open the window wide enough for me to kind of crawl through - and then I'd trust what I've always trusted before which was the bibliographies of published works' [Sociology 11D 369-410]

This also underlines the close relationship that can exist between starting and chaining. Many of those interviewed expressed the view that they would very rapidly start chaining from references they had obtained by other means.

> 'You start off with some particular piece of work that seems to be very interesting and forthright on the issue and you acquaint yourself with that and you look at their references' [Sociology 14C 064-072]

> 'I had to start with the textbooks - some of the textbooks - in the area. From the textbooks work back to some of the references they had cited and build out in that way' [Economics 1C 2298-2315]

> 'I started very simply I suppose, with an attack on the University Library's own index systems...in the subject indexes, and looked for books initially. Got out some, read them, started to chase these further - well chasing references from these books into more detailed material by the usual tree system of searching' [Research Student Prehistory and

Archaeology 15D 068]

Chaining

Backward chaining again represented one of the most significant features of the information seeking patterns of the social scientists interviewed. All of those interviewed employed it and for some it was the prime way in which they gathered or sought information. Sometimes this took the form of using lists of references in material as select bibliographies.

> 'The main thing in fact I've found is following up trails. You get one book which is relevant and you chase up the references which are relevant' [Sociology 10C 421]

'Through bibliographies in new work that's

coming up' [Sociology 12C 096]

'It's quite often a question of working your way back...picking up things which have been published two or three years ago - as you see them cited elsewhere' [Economics 1C 216]

'Certainly trying to follow up references that you find in the literature that you come across' [Sociology 14C 131]

'There are textbooks that keep appearing, and what I do is, obviously, read through these and keep as up to date as possible, and also go through the bibliographies at the end' [Economics 3D 057]

'It's a matter partly of you get some references and you follow them back and then find other references in those same articles' [Politics 9C 238]

It was interesting that a number of the social scientists from all the departments had employed forward chaining at some time or other.

'When I started on the initial piece of work I'm doing I spent several days with the Social Science Citation Index and particularly following two or three key references back...One of the references I had...was 1972 - 1973 so I was looking for citations of this particular reference, and then as I got some more modern ones coming up

I would read these, look at their reference lists and work back from them as well, to look back at the articles they had in turn cited in addition to my key one' [Economics 1C 348-403]

'We...used the citation index to really see how that had spread out, who else had used it and so on' [Sociology 13C 493 671]

'I didn't just take the online search as the beginning and end of it - I then chased everything I could from that, both up and down' [Research Student Prehistory and Archaeology 15D 210]

From the interviews with the researchers at the MRC/SSRC Unit, the psychologists, and the educationalists, it was clear that the extent to which chains of citations will be followed depends on a number of subjective factors relating to the perceptions of the person doing the searching, as does the decision when to finish chaining with some degree of confidence that the important lines of work have been covered. And, as one sociologist remarked, the activity itself can be pursued with different degrees of intensity.

'It depends very much what stage you're at. If you're at the initial exploring stage you generally chase quite hard to make sure that you have got the relevant stuff. If you're at the stage of writing up and you're writing like a madman and you want to try and make

> sure you've covered the really important references the moment you think you've covered your rear you stop - because of the problem of the deadline' [Sociology 10C 427-428]

Another feature of chaining mentioned, which had also been remarked on by the other groups, was the extent to which they were, or would expect to be, familiar with references in areas they were well acquainted with. The extent to which the social scientist would expect to be familiar with references was directly related to the amount of time which the person had worked on the particular subject, although new aspects of work would tend to lead into areas where the individual would not expect to be so familiar with the material cited. Again, as with the other groups, seeing the same, or corroborating, or increasingly marginal references was a factor in the decision to bring chaining to a halt.

> 'There are certain areas which I think I'm quite well acquainted with the history of the literature, so there will be many cases where the references cited at the end of them are familiar to me, or the majority are familiar. I don't always backtrack on them, it depends if there's a particular point or particular purpose or me reading that article. My current area of interest is in particular aspects of family buying behaviour, and I'm still coming across material there, sometimes from the 1960's and 1970's, which I'm not familiar with so I'm doing a bit of backtracking there, by browsing...also by picking up the back articles, and that's generally from the references that are cited' [Economics 1C 348-403]

'Eventually you find, I think, that either you feel you have actually mastered the field, in the sense that books continually throw up bibliographies with which you are familiar, and the number of references that you haven't heard of get fewer and fewer, and sound more and more marginal. Either that happens or you end up thinking, I have done enough background reading now that to do any more would simply stop me from doing what I've got to do, which is to get on and think for myself - and so you decide to pack up anyway' [Sociology 11D 369-410]

Browsing

Several of those interviewed mentioned browsing through journals, books, or abstracts when they visited the library, or bookshops. This activity tended to merge into monitoring so that, at the boundaries, it can be difficult to distinguish between browsing for current awareness and monitoring, particularly, again, if a current contents type service is employed where

the individual can glance at the contents pages of a relatively large number of journals quite quickly.

> 'One checks out publications from the Sports Council, and mosies around the library and bookshops and so on...Obviously one comes across things by happenstance a lot of the time...I'd go to see what local libraries had in the area - as I say with the sociology of sport stuff I'd look at what our library has, what the Polytechnic library has, what the city library has, browse around under sociology, under sport, look at physical education, these three sorts of areas' [Sociology 14C 127 319-333 486-490]

> 'I look through each journal. I would see what the topics are on the the title page and if there's one which I think might be of interest I would turn to that...Go on the title first and then flick to the actual article in the journal, have a look at the abstract to see what it seems to be covering, glance through' [Geography 7C 146-158]

> 'Every now and again I go down to the library and look through the index sheets of all the sociology journals and the social welfare journals. And I'll go to the Law Library and look at the British Journal of Criminology there, because in the back of that there's a list of recent publications on the police. That has a digest at the back of it of recently published work - so I'll look at that for articles...Every time I go to the library I tend to look - I just browse through to see if there's anything of interest' [Sociology 12C 124-140]

> 'From time to time I've leafed through Economic Abstracts' [Economics 2D 591]

The role of differentiating in distinguishing between browsing for current awareness and monitoring is clear from the comments of one of the economists interviewed.

> 'Typically by using Current Contents in Management and Contents of Recent Economic Journals, and they cover most of them - most of the contents of the journals...They are held in the coffee room. When one goes in there to make a cup of coffee or tea I flick through the series and look through the last lists. If there's anything or a series of three or four which I spot and I particularly want I make a note...There are a lot of irrelevant ones included which one doesn't have to spend any time with, so some are bypassed such as the accountancy journals. Some, the Journal of Behavioural Economics for example, it's a journal which might carry one or two articles a year which I'm interested in and want to follow up, so that's a cursory flick. The others, the more central ones - the Journal of Marketing; the European Journal of Marketing; the Journal of Consumer Research - then obviously one gives more attention to' [Economics 1C 258-292]

Differentiating between sources in terms of the likelihood of their containing material on the substantive topic, from a particular approach or perspective, or of the right quality or level was a feature of the information seeking patterns of many of those interviewed. Particular journals were often identified as being likely to contain most of the material pertinent to them.

> 'I suppose primarily by seeing the relevant specialist journals, which are Imago Mundi, it's an international journal but it's published in this country, and the other one is The Map Collector. Plus various cartographic journals that contain occasional historical material articles in them' [Geography 6C 031-050]

> 'I have most of the journals I use. All the main ones that I need or use regularly. I've subscribed to many over the years, it's just a question of pruning it down to the ones that are the most useful. I've confined it to half a dozen now that I find that use more regularly: Survey - a Soviet east european studies type journal - Problems of Communism, Soviet Studies. I suppose these are the three main ones. And then a lot of Soviet newspapers' [Politics 8C 165-174]

> 'On the long wave I would mention two or three books and one journal in particular - Futures. It's supposed to be a journal for planners and forecasters, so it tries to take a sort of future oriented approach on quite a number of social and economic issues, and it's interested in the long wave because any insights into the long wave can give you insights into the future. So there are many, many, more articles in Futures about the long wave than just about anywhere else - although Review has got a few' [Economics 2D 626-655]

The differentiation of material resulting from different approaches to a subject was particularly marked for the research student in prehistory and archaeology. In his case, two different research communities were working on subjects which related to different aspects of his problem, but from different directions and with different interests, they published in entirely different sets of journals, and, were either completely unaware or uninterested in each others work.

> 'Between literature sets, in this case...They never cite each other. They are obviously unaware of each other...90% of what I'm talking about is journals...They were quite separate journals completely. I mean we're talking about different epistemic communities essentially. They happened to be both working on a problem that generated information which was suitable for my research' [Research Student Prehistory and Archaeology 15D 177-204]

Differentiating sources by the level of treatment of subjects was a common feature of the social scientists' information seeking patterns. Sometimes

this was employed to filter out material which was likely to be too specialised or technical for the purposes of the individual carrying out the search. The individual was only likely to consult such material if it was critical to an understanding of the issues he was addressing, at the level he was addressing them.

> 'Well, there are the law and accountancy journals, and on first principles I wouldn't look at them unless I saw a reference to a specific article...I know that some of the journals that I could look at would produce material that was incomprehensible both to me and to my students, although the message of the article might be quite relevant' [Economic and Social History 5C 109-119]

> 'There are say thirty-five people involved in some technical problem, in the country, and they get - these thirty-five people - four articles a year in the Economic Journal and two in Economica or whatever...They are a group of people who are genuinely addressing each other on what they see as a technical problem, but they are very largely technical issues, analytical technical issues, which are quite genuinely being addressed...but it is the specialist addressing the specialist, largely at a technical level, rather than someone addressing a more general problem for a more general readership...I flick through them to make sure that they are not doing anything different - to put it crudely - that they are not widening into issues which have any sort of immediate relevance for policy' [Economic and Social History 4C 139-158]

> 'I keep an eye on the specialist journals, but more generally I go to the specialist journals when I find a particular reference from somewhere else which refers to the

particular article. I tend to do it that way around' [Politics 9C 190]

A distinction which had been made by the earlier groups between 'academic' and 'practitioner' type journals was again made by two of the sociologists interviewed.

> 'The British Journal of Social Work is the academic journal of social work in Britain, and it takes articles mostly by academics about the entire field of social work...The The Journal of Adoption and Fostering is specific to that field of adoption and fostering only. It has much shorter articles, often written by practitioners, or at least by people with one foot heavily in practice, it's much more up-to-date. The British Journal of Social Work suffers like all other journals with this kind of eighteen month time lag from submission to publication very often. Adoption and Fostering doesn't and Adoption and Fostering does take a responsibility for keeping you up to date in certain areas in the field...I read Community Care, and Social Work Today - they are chatty and newsy things - as well' [Sociology 13C 537 574]

> 'I think in the trade journals is one place. By that I mean things like Community Care, Social Work Today. Or the heavier ones like the British Journal of Social Work or The Journal of Social Policy. They are thicker, they have more solid articles, they are not popular. Community Care is often described as the Daily Mirror of the social work world' [Sociology 10C 059-069]

A further way of differentiating sources in terms of their level of treatment, which was mentioned by

both the economic and social historians interviewed, was in terms of the form in which the material was published.

> 'As far as one's research is concerned I would be more interested in the monograph stuff - journal articles would be the more relevant but publishers are still producing monographs - therefore books are still useful. As far as teaching is concerned then it is usually the book which is more relevant - text books obviously - some of the less technical literature. Although the dividing line between something that you're interested in for teaching and something that you're interested in for research is very, very difficult' [Economic and Social History 5C 280-297]

> 'It's usually in specialist tomes - you know they're large volumes, they're big books - where somebody is doing quite a large piece of work...and every time a new one comes out, then, obviously, it's lionised for a time and everybody knows about it. And then the fashion wanes and one gets a whole series of pamphleteering...There isn't all that much in the academic journals...the hard academic journals, because they are more concerned with the technical aspects than with policy related aspects. So what you get is a series of second rank journals that deal with political economy and also a lot of pamphlets. The Institute of Economic Affairs, this sort of thing, and then there's a sort of war of pamphlets' [Economic and Social History 4C 062-083]

And, again it was the experience of some that differentiating a limited range of sources on which they could concentrate their attention was not always either practicable or easy.

'The Journal of Consumer Research...the Journal of Advertising Research, which is another fairly core American journal...the Journal of Retailing, which sounds very trade but is an old established journal from New York University...the Journal of Consumer Policy...the Journal of Consumer Affairs...They were probably them all...They're fairly central to my own area...There are a lot of other journals - the place is stacked with journals, the whole discipline is riddled with them. It's a particular problem that one has: you can't easily identify four or five very core journals and follow them and pick up the rest from elsewhere. Your material might be scattered over fifteen to twenty management journals, all of which are capable of carrying, from time to time, very, very, relevant articles ' [Economics 1C 182-195 222-231]

'Stuff does come up which is relevant to my research interests in any number of journals because I've put myself, intellectually or academically, on the boundaries of several subjects. One is looking at a wide range of stuff. Basically what I do is go along to the libraries and look at the current copies. I go along the shelves every now and again picking out those which I would think are most likely to have stuff which is of teaching and research interest, from past experience, I would think that there must be at least twenty economic journals alone, then there would be a few sociology journals, the number of journals in economic and social history is quite small and they're ones that I take' [Economic and Social History 5C 084-105]

Monitoring

Informal contacts represented a significant means by which some of those interviewed kept up-to-date with developments in their areas.

> 'Being in on the network - people who are working in this field, most of whom I now know personally. They have written some of the books - but we will pass around information. We meet from time to time at conferences about two or three times a year...A very large part of the important information I get from actually meeting people working in the field...Personal contacts play a very large part in our collection of information. I'm not someone who devours prodigious numbers of books and magazines - so I rely on that quite heavily...I should be most surprised if there was a large area of work which we are simply not aware of. We know quite a lot of other people working in the area and we actually send working papers out to other people who then can say - ah, but you have not mentioned so and so - so that's another way of checking' [Sociology 10C 094-101 112-115 439-445]

> 'I do now know a lot of people spread about who are working on similar themes to mine. About twenty to thirty people that I'm in reasonably regular contact with. Sometimes exchange of papers or draft manuscripts for commenting on. Participation in the specialist groups, going to give papers at their institutions, more casual contact by phone' [Politics 9C 515-540]

'I think it's mainly that one knows what research is being done in the field and by whom, and one knows what is likely to be published. And, if one is aware of a gap, an information gap, then one can very quickly and easily get to the filling of that. Either through personal contact, or by looking in a research directory, or by looking at the journals. The journals would probably be about the third resort - personal contact and the research directories probably more important' [Geography 6C 093-104 211-225]

One lecturer was particularly dependent on contacts made while working as a research student. From these the person was able to employ other individuals and organisations to act as a monitoring service for him.

'Merely by meeting people in the course of initial research work as a student and post-graduate student and being over in Canada and America...It is a sufficiently small and compact field for us all to know each other more or less' [Politics 8C 061 555]

'We have a number of organisations within organisations dealing with these topics now. For example, there is an association for the study of nationalities in the Soviet Union. So that is a major source. It produces a newsletter and a bibliography...These tend to bypass either the major journals or the major bibliographical material that one would normally use probably...Again it's a question of private contacts rather than any formal structure' [Politics 8C 040-048 070-081]

'I have contacts mainly in Italy where there's a large Albanian community and again

> have been able to follow it up through that. So I receive all their publications, as well as summaries of - abstracts - and sometimes the original materials from Albania as well. Again it's a question of private contacts rather than any formal structure...They are really a kind of filter system...In the sense that, for instance, with the Ukrainian press there are obviously several hundred newspapers and, or, government publications coming out of that one area. The Ukrainians themselves, the emigres, produce a thing called The Digest of the Ukrainian Press: that's a regular publication, in America in that particular case - but what it does is simply translate major articles that have appeared that week or that month in the press in the Ukraine. So whereas I might, on a regular basis, look at one newspaper, this will monitor perhaps a hundred and check out some of the more important items, and they also do a service which is non-translated - simply photocopies of the originals - and there are a number of outfits and organisations that do this. So the primary materials - I don't have to have the entire bulk of them coming in and they are filtered a little for me by the various groups and organisations' [Politics 8C 088-114]

This reliance on others to monitor the field was also mentioned by one of the sociologists interviewed.

> 'One checks out publications from the Sports Council...Because they themselves have commissioned a fair amount of research as well...I've got a colleague...who gets stuff from them, so I rely on him at the moment to get the material through' [Sociology 14C 127 143 146]

Monitoring the journal literature was mentioned by several of those interviewed, this usually involved prior differentiation of a limited number of journals to monitor.

'Its a matter of keeping up to date with the subject through the journals I suppose primarily...I don't read too many books on it...There are five or six major journals which I use regularly. I make sure I read these when they come out...Obviously having done a degree in Russian Studies I know the sort of materials that you can use to study such a phenomenon...So I have some obvious awareness of what the major journals, official publications, and books are on that area. So I'd know what to go to almost immediately - I'd have no difficulty with that' [Politics 8C 232 240 441-454]

'I suppose primarily by seeing the relevant specialist journals. Which are Imago Mundi, it's an international journal but it's published in this country, and the other one is the Map Collector. Plus various other cartographic journals that contain occasional historical articles...I think I've always known - not quite that one was born with that information - I suppose that as a geographer with an interest in historical geography, which I have always had, I've known of Imago Mundi for most of my professional life' [Geography 6C 031 093-104]

'Looking through the current journals as they arrive in the library and they're obviously kept on the shelves for so many months before binding - so I allocate that amount of time for looking through skimming various articles and taking notes...The other thing I do, as far as keeping up to date is concerned...is to spend every Friday afternoon once a month, say, looking through the current journals as they arrive in the library' [Economics 1C 258-281 321]

'The British Association of Adoption and Fostering...have a journal called Adoption and Fostering, which is really very good, and

is very up-to-date, and does have the main issues, and does publish research. It has very good book reviews, and keeps you up to date with issues and literature in the field, and I certainly use that - I'm a personal member of the British Association of Adoption and Fostering - you can go back through it and check out references, and themes, and book reviews, and abstracts, and so on, within it' [Sociology 13C 167-187]

Two of those interviewed used the press in a monitoring role. In both cases their research interests were connected with topical or policy issues, and the press was used as the basis for keeping in touch with what were the current themes of the debates.

'The heavy press - the Financial Times - which is what I use for most of them - give me a one or two page summary. I just cut the Financial Times summaries - in the files here I have twenty odd years worth of summaries of all the forecasts - and if what they say looks different or new or seems to have more data in, as summarised by the press - I read the Times and Financial Times summaries - then I'll get the thing from the Library and have a photocopy of the relevant sections' [Economic and Social History 4C 097-103]

'There is a lot of commentary all the time in the press. Some of it a very high level in the quality papers - even the less quality papers...There is some interesting material - particularly when major issues come up. It's interesting to see the range of opinion which exists within the Party. The range of viewpoints which are put forward into public debate which are recognisably within one political party and directed at influencing opinion within that party...It's particularly true within the Conservative Party because so much of the press is oriented towards the Party...I try and read as much of the press

> as I can - The Financial Times, The Times, the Guardian - as well as some of the journals - the Economist, the more popular economic journals...Also through the pamphlet literature. I keep abreast through the pamphlet literature, I think that's very important. Partly through regular trips to certain bookshops which stock this stuff. Partly by being on the mailing list of various research institutes and asking for their latest lists of literature. Those are the main ways' [Politics 9C 059-067 190-213]

Only one of the social scientists interviewed mentioned making particular use of a current contents type service: these were Contents Pages in Management, and Contents of Current Economic Journals. This lack of use of Current Contents was similar to the educationalists, but contrasted with the use of the service by some of the researchers at the MRC/SSRC Unit and by many of the psychologists.

> 'It's a matter of monitoring the very poor abstracting services that exist in the area...and two things called the Contents Pages in Management, and Contents of Current Economic Journals. The Journal of Consumer Affairs is on the fringe of these two and sometimes they seem to include it, and sometimes they don't...They cover most of the contents of the journals, and they are published on a fortnightly basis' [Economics 1C 202-214 258-281 292]

Again in contrast to the researchers at the MRC/SSRC Unit and psychologists, and again similar to the educationalists, several of the other social scientists interviewed made reference to using publishers' lists to keep-up-date. Book reviews were

also mentioned as useful in this context.

> 'You have the usual grapevine things. You have publishers' sales literature, which comes round with advance notices...Reviews in the journals, in the things like the Financial Times and the Times' [Economic and Social History 4C 168]

> 'I come across stuff in reviews. Publishers' literature is always very useful' [Economic and Social History 5C 234]

> 'Keeping tabs on books as they are published. That's a very mixed business - it's partly publishers sending information about their books. It's partly getting information from book reviews from the journals I mentioned earlier plus various geography journals. It's partly through the personal information network other people working in the field. Mostly I find I know what is going to come out because I know the people who are producing it' [Geography 6C 031-050]

> 'In part through the grapevine of other scholars that I'm in contact with. In part through book reviews and publishers' advertising...Through bibliographies in new work that's coming up. Through the academic journals themselves...One finds articles there - but they often trigger off a book that's going to be written, published after, this is the first piece to a book. The publishers themselves - through the central computerised index of scholars interests...I just put 'criminology' down there' [Sociology 12C 096-112]

Extracting

Extracting from key journals was mentioned as an effective means of quickly gathering pertinent material, either through using indexes to the journals or by actually working through the contents pages of individual issues.

> 'Devonshire Archaeological and Exploration Society is actually held in the local library and one day I simply - or two days three days whatever it is - I simply worked my way through the whole bloody thing, chabong, literally parried my way though it because its not that big. I went through all the recent issues that way - I think that have an index in that too and I used that - but all the more recent issues I actually went through, thumbing through to see - literally because I felt that the indexing of these sorts of things is not that hot, and I wanted to look through and see what was there in totality' [Student Prehistory and Archaeology 15D 622-634]

> 'I started with the last five years Economic History Reviews - which are standard - well you know that is the economic historian's 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...Then there are 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 456-473]

'It's a matter of painstaking slog. Of just going through indexes of a journal like the American Economic Review and so on, and just picking out articles that might be relevant - that look interesting...Going through the indexes going back normally as far as the journal goes back' [Politics 9C 240-254]

'I look through each journal. I would see what the topics are on the title page and if there is one which I think might be of interest I would turn to it' [Geography 7C 335]

The similarity of form between monitoring and extracting was clear. One mentioned that extracting from sources was subsequently complemented by monitoring those same sources. Several others employed extracting as a means of monitoring, with different sources being the subject of separate extracting exercises, sometimes taking the form of a rolling programme carried out over a number of years.

'I went through the indexes of of what was in the library - the journal indexes - cumulative journal indexes exist in most of those. I searched through all of those...And subsequently, those journals, all those ones that were present that were relevant that had been index searched, and so on, I kept up to date with by regularly scanning stuff as it came in. Scanning each new issue, quarterly

issue, bi-monthly, checking the title of articles on the front covers or back covers or inside pages to see if any of the titles appeared to be relevant' [Research Student Prehistory and Archaeology 15D 243-262]

'Over the years I suppose I've been through all the county archaeology society journals that have ever been published. I've had a sort of consistent programme of each year of working through every volume of two or three counties' [Politics 8C 240]

'I look through journals in some kind of order...I normally do a comprehensive search of a certain journal. I start at the latest edition and work backwards and try and get as much up-to-date material as I possibly can' [Economics 3D 038-043]

'I have a system whereby I have a list of journals that I normally look at, and I make a note of the last one that I looked at...And so I work through journals and I have a list of what was the last one I that I referred to, and so - usually over the summer or winter vacations I have a session in the library to read through to see what papers have come out' [Geography 7C 098]

Extracting from publishers' catalogues was also mentioned as a significant activity by a social scientists from almost all the groups. Many obtained these lists through the services of a computerised mailing service.

'I'm a very consistent reader of publishers' lists - I go through them' [Politics 8C 190-195]

'I read publishers' catalogues. They send them. You get a little form comes round every year asking you what your interests are. I tick the various areas. I tick industrial sociology, I tick organisational behaviour, and so forth, and, in due course along comes these vast shelves of bumph that I work my way through usually, and find books that I then track down' [Economics 2D 161]

'I do also make a point of going through the publishers' blurb that comes through - I don't throw them all away, I do go through them, very quickly' [Sociology 10C 083-086]

'IBIS (International Book Information Service)...it's a computerised guide, every year they send you a sheet of broad subject areas and narrow subject areas. You tick those in which you have a teaching or research interest and which you want to receive publishers' information. It's a centralised mailing service. It's very good you get an awful lot of mailing. It does enable you to know in broad terms what is being published. It seems to be fairly comprehensive of the major academic publishers at any rate' [Economic and Social History 5C 234]

'I get notified of their (publishers') catalogues and lists and new publications through IBIS - International Book Information Service' [Sociology 11D 220]

Two of those interviewed mentioned extracting from bibliographies as useful.

> 'There was an excellent bibliography of Dartmoor...A pretty total one...It had been done, the original, only about eight or ten years before and he'd updated it since...And I went through that, just read it cover to cover actually. I went through every subject, I went through the whole thing very carefully indeed, making up index cards to any articles that seemed to be relevant' [Student Prehistory and Archaeology 15D 640-692]

> 'What I tend to do also is go through the annual bibliography as well - thats one thing I do use in the Library - British National Bibliography - I find all the new books published in English on Russia' [Politics 8C 190-195]

Conclusion

This group represents the most disparate of those described in terms of the disciplinary background, research and teaching interests, and departmental affiliation. The group was similar to the psychologists and educationalists in that - with the exception of the research student interviewed - all the social scientists were engaged part-time in teaching and part-time in research, and so had to balance the demands of the one against the demands of the other. The research interests of the other social scientists were similar to those of the educationalists, in that they tended to be specific to the individual concerned and there was little scope for collaboration on research within their own departments. Similarly, the induction of research students was relatively sporadic.

In terms of their information seeking activities there were both similarities and differences between this group and those of the others described. Some had made use of forward chaining, which was similar to the researchers in the MRC/SSRC Unit and to the psychologists, but different from the educationalists. However, similar to the educationalists, but different

from the researchers in the MRC/SSRC Unit and the psychologists, there was virtually no use of Current Contents. Publishers' lists were mentioned as a significant source of information by some of the social scientists, which was again similar to the educationalists but different from the researchers in the MRC/SSRC Unit and the psychologists. Some of the sociologists also stressed a distinction which had been made by some of the educationalists between academic and practitioner journals.

Overall, the information seeking patterns of the other social scientists could be accommodated within the framework of the behavioural model without modification of the form or content of that model as it applied to the other groups.

Chapter Eight

A Behavioural Approach to the Design of Information Retrieval Systems for Academic Social Scientists

The overview of the features of the behavioural model and variations in detail of this model within and between the broad subject groups of the social scientists interviewed have been set out in the previous chapters. From this analysis it would appear that the model is sufficiently flexible to reflect the

information seeking activities of academic social scientists across a broad range of social science subject areas. In this chapter it is proposed to employ the features of the behavioural model to derive recommendations for information retrieval system design, and to relate these recommendations to other work undertaken on human-computer interaction. In addition, some operational and experimental systems will be examined to assess the extent to which the features of the model are present in existing systems, and the requirements for implementing the behavioural model on an operational system will be outlined.

Human Factors and Information Retrieval Research

Research on human-computer interaction has included a substantial body of work relating to computer-based information retrieval systems and services. There have been a number of reviews of different aspects of this work, covering studies of

interface design for interactive information retrieval (Thompson, 1971), the user interface in interactive systems (Bennett, 1972), human factors and information retrieval (Martin, 1980) user computer interface designs (Brenner, et al., 1981), the psychology of human-computer interaction (Moran, 1981), recommendations on the design of man-computer dialogues (Maguire, 1982), human factors in interactive computer dialogue (Ramsey and Grimes, 1983), and human-computer interaction in information systems (Belkin and Vickery, 1985).

Nevertheless, despite the interest in this aspect of human-computer interaction, there has been little in the way of real improvement in system design (Belkin and Vickery, 1985), despite the fact that human factors research played an important role in the development and implementation of computer-based information services as Martin (1980) pointed out -

> 'In the development of existing information retrieval systems, human factors considerations have played a role' (Martin, 1980:171)

Martin cites, as examples, Katter's (1970) review of user acceptance factors undertaken before designing the ORBIT information retrieval system, the pretesting of Lockheed's DIALOG system at Stanford and its modification in response to the reactions of the users (Timbie and Coombs, 1969: Summit, 1971).

Martin (1980) also noted other studies which had either provided guidance as to how user reaction might be employed in the design of computer-based information retrieval systems, or had been considered in the design of operational systems (Marcus, et al., 1971: Martin and Parker, 1971: Thompson, 1971: McAllister and Bell, 1971: Walker, 1971). A review of the work undertaken during that period on the factors and considerations taken into account in designing the human-computer interface for the interactive computer-based information systems then being developed is given by Bennett (1972). Martin (1980) concluded that

> 'Since that time the human factors contribution has been present but has not materially altered system design' (Martin, 1980:172)

Belkin and Vickery (1985) reported a similar position in respect of any changes in system design.

In fact, outside the area of interface design the contribution of human factors research to the design of computer-based information retrieval systems has been negligible. Moreover, study of human factors in relation to the fundamental questions of database design and construction, as opposed to the study of human factors in the design of interfaces to databases, has not been undertaken. Possibly this is due to the human factors researchers misunderstanding the nature of the problem underlying human-computer interaction in

the context of computer-based information retrieval systems. Martin drew attention to this misunderstanding when he noted that

> 'When retrieval systems were being developed, it was assumed that that end users would search for themselves. This has not happened in practice...Perhaps one reason why human-computer researchers and designers felt that end users would search for themselves was they they tended to view the task as one of human-computer communication. Instead the task is one of utilizing data bases and the searching aids associated with those databases' (Martin, 1980:166-167)

Developing a comprehensive picture of human-computer interaction in the context of computer-based information retrieval systems requires an understanding of both the nature of the interaction with the database and the nature of the interaction with the interface to the database.

A result of this is that a considerable amount of research has been addressed to issues connected with interface design (Bennett, 1972: Martin, et al., 1973: Brenner et al., 1981) and this can be used to provide guidance on the features which might be desirable in an interface for a computer-based retrieval system. There is also guidance available on the general system features desirable for optimising the human-computer dialogue (Maguire, 1982: Grimes, 1983). On the other hand, there is very little research which directly relates to issues concerning the design and

construction of the database of a computer-based information retrieval system. The two aspects of database and interface design are related and in the treatment which follows questions of database and interface design will be treated together.

The Application of the Behavioural Model

The behavioural model derived from the analysis of the information seeking patterns of the social scientists can be employed in two ways in the context of information retrieval research -

1. as an evaluatory tool to identify the existence and ease of implementation of features of the model in existing systems;

2. as a basis for design recommendations for future systems.

It is proposed here to concentrate on the second aspect of use, and to employ the behavioural model to analyse the requirements of an exploratory information retrieval system for academic social scientists. A brief consideration of the former aspect will be provided in the context of an examination of how an operational system based on the recommendations outlined here might be implemented in practice.

Starting

The significance of starter references - key articles from which the searcher can proceed to

identify other material by chaining - has already been pointed out. 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. 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, then the system is likely to be used in an attempt to identify material which seems potentially useful in this role. 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. The references which appear interesting to the searcher may be noted, and, again, the searcher may wish to proceed straight to chaining, or some other information seeking activity.

Starter references are usually employed in two roles -

1. to alert the individual to principal ideas or key studies;

2. and to provide overviews of an area and serve as the basis for chaining.

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 these two roles. 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.

Chaining

Chaining is a principal feature of the information seeking activities of social scientists. As already mentioned it has two aspects - backward chaining and forward chaining. Backward chaining represents a major way in which many social scientists locate material of interest, and forward chaining, though less commonly employed is also a significant way in which some social scientists pursue information. A retrieval system designed for academic social scientists would need to have 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 to 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.

Browsing

Browsing, in the context of the behavioural model, is understood as constituting semi-directed searching in an area of potential interest. The simplest way to incorporate a browsing facility into the system is to

allow the searcher to browse, while the system provides the information the searcher requires to browse effectively. If the searcher does the browsing, not the system, then the cognitive aspect of *browsing does not* need to be specified. In this sense, providing features which allow browsing is relatively simple; far simpler than devising a system which builds up a picture of the searcher's cognitive world in order to do the *browsing* for him. Providing facilities for the searcher to browse through the database solves the same problem without the creation of a complex representation in the system of the searcher's cognitive world.

A system designed around the concept of exploratory capability should permit 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, and subject terms and broad subject headings.

Providing access for browsing for most of these types of information is relatively unproblematic, for most of them 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. 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.

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 needs to be set up in such a way that browsing can take place easily within and between the different categories of information. This requires that the database be constructed along lines that provides browseable lists of these various categories of information, and a simple means of moving between lists and obtaining further information on items in the list. As with the other facilities the means of identifying material should be simple, and the searcher should be aware of the different browsing features available perhaps again through top level screen menus.

The browsing facility might be enhanced by the use of graphical displays of terms. 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.

...ifferentiating

Differentiating between sources in terms of their likelihood of containing material of interest was frequently mentioned by the social scientists. Three principal differentiating criteria were employed - differentiating by substantive topic; by approach or perspective; and by level quality or type of treatment. These three criteria represent different types of filter on material examined, but they are all effected in the same way, by the social scientist identifying certain sources or types of source as being more likely to contain material of interest. Types of source could include journals, conference proceedings, monographs, research reports, theses, or newspapers. Other differentiating criteria which could be employed might include the language, country, or date of publication, and the the number of pages or references.

Conventional computer-based searches, 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 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 creates 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 were 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

onto 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 easily differentiate 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.

1. To restrict the search to a limited set of sources or types of source;

2. to exclude certain sources or types of source from the search;

3. or 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. The searcher should be able to carry out a comprehensive search of the full range of sources at one time - particularly if there is little material and what there is is widely scattered. However, if the searcher does wish to restrict the search to specified sources or types of source then this option should be available. The important thing is to give the searcher the choice of options, not preclude them.

Monitoring

Implementing a monitoring facility in an exploratory 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. In this way, the monitoring facility may 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.

Extracting

From the interviews with the social scientists extracting describes how searchers either worked through individual runs of journals, sets of publishers' lists, bibliographies, indexes or abstracts, or consulted cumulative indexes to such sources. Both types of approach to extracting should be provided in a retrieval system as some searchers prefer examining the contents of individual issues themselves, 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. Other searchers prefer the convenience of employing a cumulative index to sources. To reflect these different preferences facilities for both types of approach to extracting should be provided in the retrieval system.

For the first approach to be possible the system must be designed to allow continuous movement through different source streams. Whether these source streams be runs of journals, publishers' lists, sets of conference proceedings, laboratory memoranda, working papers, or whatever other form of material is held on the database. This approach to extracting requires that the sources be recomposed more or less as they were originally composed, so that the searcher can proceed through streams of these recomposed source materials.

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 to the searcher 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 allow the searcher to easily limit the search 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.

Although many of the recommended features are available on present systems their availability is usually ad hoc, reflecting the manner in which a particular file has been produced, rather than representing a thought-out design strategy. Because they are features of particular files, not of the systems themselves, they are not integrated with each other, and are usually cumbersome, difficult to use and not well presented.

It was clear from the interviews with those social scientists 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. Apart from this there is no real equivalent of the more complex starting facilities indicated above in conventional computer based information retrieval systems and services.

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, as cited references are identified by prefix codes in phrase indexed fields of the record, no abstracts are available, and the only form of subject access is via the terms occuring in the titles. In consequence, as Morehead and Rouse (1982) observe,

> 'no readily available database provides a user with online access to articles, abstracts, text and reference lists' (Morehead and Rouse, 1982:196)

In the absence of an operational database which possesses these features tests on 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') (Morehead and Rouse, 1982: 1983: 1985a: 1985b: Rouse and Rouse 1984: Rouse, et al., 1982).

Using the experimental 'DBASE' system enabled close examination of the human computer interaction involved in searching that system, but the limits of that system place constraints on the type of searcher behaviour it is possible to simulate. In fact, the only types of searcher behaviour it is possible to study in the 'DBASE' environment is '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 (Morehead and Rouse, 1982).

In conventional retrieval systems facilities for browsing are virtually non-existent. 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 (Hildreth, 1984). 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 databases, DIALOG using the EXPAND command, BRS employing the ROOT command, and ORBIT using the NEIGHBOUR command (Schuman, 1983).

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 (Ayris, 1985). Davies (1978) devised a mechanism for browsing based on term frequency ordering of documents which starts from an unstructured search consisting of a few keywords, the search being subsequently, interactively, refined in response to user reactions to documents retrieved. Another 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 (Hills et al., 1983).

Visual approaches to browsing have been put forward by Frei and Jauslin (1983) using a graphical approach, and by Palay and Fox (1981) 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 (Bertrand-Gestaldy and Davidson, 1986), and an experimental graphics package for information retrieval has been developed by Sanderson (1984) which could be implemented on an operational system. A similar type of approach is embodied in the experimental PRIDE retrieval system developed by Bovey and Brown (1987) 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 (Curtis et al., 1981). The BROWSER terminal system has also been employed in the Public Information in Rural Areas Technology Experiment (PIRATE) where a database management system containing tourist and public information is operated by simple screen displays on the 'touch sensitive' screen.

Perhaps the most interesting operational application of the browsing concept in database design has been that employed in the development of IME's TinMan software, where the searcher is able to 'Browse' and 'Navigate', following up connections between items on the database (Noerr, 1979: Noerr and Noerr, 1985). 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' (Ayris, 1985:87-88)

The underlying design philosophy of this approach is similar to that underlying the general recommendations for an information retrieval system based on the behavioural model, which is very evident in the associated Hypercatalog proposals (Hjerppe, 1985).

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 be 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, as the sources and types of documents are identified by prefix codes in phrase indexed fields of the record. 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.

Facilities for monitoring in conventional computer-based services are generally restricted to current awareness searching, although theoretically it would be possible to reconstruct individual issues of journals for contents page monitoring by careful application of the same features just described. Similar restrictions apply to extracting. In particular, there is no provision in existing computer-based systems for searchers to proceed through contents pages of sets of journals.

It is worth noting, also, that although the form of extracting which proceeds through the use of cumulative indexes may appear identical to a conventional computer-based search, it is, in fact, different in kind. The searcher manually scanning a cumulative index is employing interpretive skills and working with the meanings of the index descriptions, the computer is carrying out a matching process on

symbols ultimately only 'recognised' as binary code. The failure to recognise the difference in kind between the two types of searches, or more generally to be aware of the difference between how people think, and what the computer does, either in the context of searching a computer-based information retrieval system (Ellis, et al., 1987: Nelis, 1985), or in other forms of human-computer interaction (Pinsky, 1983), is the source of many confusions.

Implementing the Behavioural Model in an Operational System

The recommendations outlined do not, in themselves, represent a model of an exploratory information retrieval system for academic social scientists: they indicate features that such a system should possess. In practice, a system which possessed

the facilities outlined would need to be set up so that information on the database could be presented in as many ways and forms as seemed natural to the person using the system. This means that information on the database needs to be held in such a way as to permit searching to be carried out in the variety of ways in which users might actually wish to explore the database.

The user should be aware of what avenues of exploration are open to him and what is taking place when certain avenues of exploration are followed. Otherwise there is a danger that the system will be doing something entirely different from what the user intended or that the user will misinterpret what the system is doing or can do (Ellis, et al., 1987: Nelis, 1985). What this means is that the user should be given a clear idea of what the system can do, or possess 'systems transparency' (Maas, 1983).

Implementing the recommended features in an operational system involves three general considerations -

1. the selection of material to go into the database;

2. the form of entry of individual items in the database;

3. the nature of the software required to support the recommended facilities.

The first of these questions can only be answered in the context of a particular application for a particular user group. The second and third can be treated separately from any particular application, and are related to each other.

The search facilities recommended are virtually all field dependent, so the form of entry of individual items needs to be such that the different entry fields can be differentiated by the system. The system must be able to find and discriminate between fields relating to titles, authors, sources and source types, descriptors, and citations. The simplest way to discriminate between these elements of the record is to place them in separate fields. This is the underlying basis of all bibliographic retrieval systems and can be formalised in a way which can be employed by systems analysts to devise appropriate forms of file structure and access (Momenee, 1982), and also underlies the model proposed by Kircz and Bleeker (1987) for employing a relational database for electronic and conventional publishing.

The software package on which it would be easiest

to implement the basic features outlined would be on a modified version of the Tinlib package of IME's Tinman software. This has the advantage of being based on a design philosophy of flexibility of approach to retrieval through the 'Browse' and 'Navigate' facilities, and the further advantage that work has been carried employing the package in conjunction with a 'touch sensitive' screen (Bivins-Noerr and Noerr, 1982), which might represent an option worth evaluating in the context of an operational system.

Employing the TinLib variant of the Tinman software would have the added advantage of facilitating database creation through the automatic uploading, reformatting and integration of records previously downloaded from other databases. A disadvantage of the Tinman software is that, at present, it is limited to microcomputer applications which would not be suitable for a large scale system. However, a minicomputer version of TinMan is being developed and this would allow a more extensive database to be set up and tested.

In terms of implementing the model using other systems, provided that the individual entries in the database had the appropriate information in discrete fields, it should be possible to set up a system which had most of the basic features outlined by writing specific applications programs, as Sanderson (1984) did for his model graphics-based information retrieval

system. However, writing application programs is a non-trivial task, and it would be preferable to attempt to identify existing software packages, such as TinMan, on which the features of the model could be implemented.

Conclusions and Further Research

In summary, three generic approaches to the design of information retrieval systems have been outlined. That employing the information retrieval model combined with a quantitative approach to relevance judgement; that characterised as the cognitive approach; and, finally, the behavioural approach. The principal differences between these approaches, particularly in relation to the assumptions concerning the retrieval situation embodied in each of them, have been examined, and the features which might characterise an exploratory retrieval system for social scientists based on the behavioural model have been outlined.

The approach adopted here may be criticised for

assuming that the information seeking activities of the academic social scientists can be translated as a whole, and without substantial alteration, into an electronic form. In support of this it may be argued that although use of an information retrieval system may represent an aspect of the individual's information seeking activities it does not represent the whole. It should be recognised, therefore, that use of an information retrieval system, of the kind outlined here, may only ever have a rather small and specialised role to play in the total information seeking activities of the individual.

This argument has some force: it is unlikely either now, or in the conceivable future, that academic social scientists, individually or as a group, will come to rely more on computer based information services than on their other means of gathering information. And, it is unrealistic to expect that a user will abandon all his normal habits and preferences in gathering information to employ a retrieval system. But that is not an assumption that is being made here. Rather, the contention is, that if computer-based information retrieval services are to play a larger part in the information seeking activities of academic social scientists they must be designed so as to be more nearly related to the rest of those activities.

While having a search carried out on a computer-based information retrieval service involves a special

trip to the library, or other information service, an attempt to negotiate a search strategy, an apparently complex and probably ill-understood interaction with the database, and possibly a lengthy wait for the search results, then few individuals are likely to employ this as a preferred means of obtaining information. However, if the introduction of computer systems for other purposes, such as word processing or office automation, continues to progress at its present rate, and individuals become familiar with the computer systems, the use of a computer-based information retrieval system may be seen as a natural extension of the facilities of the systems.

For the present there are two ways to carry further the ideas contained in this thesis.

1. The development and evaluation of a model system based on the behavioural approach in an academic social science setting.
2. The extension and evaluation of the applicability of the behavioural model to other groups of academics or scientists.

It is proposed to pursue both these lines of inquiry in the future.

References

Adam, R. (1971), 'Social Science information and its users', **Journal of Librarianship**, 3, 150-157.

Adam, R. (1975), 'A world information system for the social sciences', **Aslib Proceedings**, 27, 286-293.

Adam, R. (1982), 'Language and information retrieval in the social sciences', **Aslib Proceedings**, 34, 394-405.

Adam, R. (1982), 'Can the transmission of sociological knowledge be made more effective', **International Social Science Journal**, 34, 329-345.

Aitchison, J. and Cleverdon C. W. (1963), **A report on a test of the index of metallurgical literature of Western Reserve University**, College of Aeronautics, Cranfield.

Allen, T. J. (1965), **Sources of ideas and their effectiveness in parallel R and D projects**, (Alfred P. Sloan School of Management: Report no. 130-65), MIT, Cambridge, Massachusetts.

Allen, T. J. (1969), 'Information needs and uses', **Annual Review of Information Science and Technology**, 4, 3-31.

Allen, T. J. (1977), **Managing the flow of technology: technology transfer and the dissemination of technological information within the R & D Organisation**, MIT, Cambridge, Massachusetts.

Allen, T. J. and Cohen, S. I. (1969), 'Information flow in research and development laboratories', **Administrative Science Quarterly**, 14, 12-19.

American Psychological Association (1963), **Report on scientific information exchange in psychology**, Volume 1, American Psychological Association, Washington D. C.

American Psychological Association (1965), **Report on scientific information exchange in psychology**, Volume 2, American Psychological Association, Washington D. C.

American Psychological Association (1969), **Report on scientific information exchange in psychology**, Volume 3, American Psychological Association, Washington D. C.

Apted, S. M. (1971), 'General purposive browsing', **Library Association Record**, 73, 228-230.

Ayris, P. (1985), **The stimulation of creativity: a review of the literature concerning the concept of browsing, 1970-1985**, (Department of Information Studies Consultancy and Research Unit (CRUS) working paper no. 5). CRUS, University of Sheffield, Sheffield.

Bates, M. J. (1979a), 'Information search tactics', **Journal of the American Society for Information Science**, 30, 280-289.

Bates, M. J. (1979b), 'Idea tactics', **Journal of the American Society for Information Science**, 30, 205-214.

Bates, M. J. (1981), 'Search techniques', **Annual Review of Information Science and Technology**, 16, 139-169.

Bath University, (1972), **Experimental information service in the social sciences 1969-1971: final report**, Bath University Library, Bath University, Bath.

Bath University, (1975), **Size growth and composition of the social science literature**, (Design of Information Systems in the Social Sciences (DISISS) working paper no. 7.) Bath University Library, Bath University, Bath.

Bath University (1980), **Toward the improvement of social science information systems: overview of research carried out 1971 - 1975**, (Design of Information Systems in the Social Sciences (DISISS) Research Report, Series A, no. 1), Bath University Library, Bath University, Bath.

Belkin, N. J. (1978), 'Progress in documentation: information concepts for information science', **Journal of Documentation**, 34, 55-85.

Belkin, N. J. (1980), 'Anomalous states of knowledge as the basis for information retrieval', **Canadian Journal of Information Science**, 5, 133-143.

Belkin, N. J. (1981), Ineffable concepts in information retrieval, in: Sparck-Jones, K. ed. (1981), **Information retrieval experiment**, 44-58, Butterworths, London.

Belkin, N. J. (1984), 'Cognitive models and information transfer', **Social Science Information Studies**, 4, 111-129.

Belkin, N. J. and Vickery, A. (1985), **Interaction in information systems: a review of research from document retrieval to knowledge based systems**, (Library and Information Research Report no. 35), British Library, Boston Spa.

Belkin, N. J., Brooks, H. and Oddy, R. N. (1979), Representing and classifying anomalous states of knowledge, in: M. and Gray, K. eds. (1979), **Proceedings of Informatics 5,** (The analysis of meaning), Proceedings of a conference held by the Aslib Informatics Group and the BCS Information Retrieval Specialist Group), 227-238, Aslib, London.

Belkin, N. J., Oddy, R. N. and Brooks, H. M. (1982a), 'ASK for information retrieval I: background and theory', **Journal of Documentation**, 38, 61-71.

Belkin, N. J., Oddy, R. N. and Brooks, H. M. (1982b), 'ASK for information retrieval I: background and theory', **Journal of Documentation**, 38, 145-164.

Belkin, N. J., Seeger, T. and Wersig, G. (1983), 'Distributed expert problem treatment as model for information system analysis and design', **Journal of Information Science**, 5, 153-167.

Bennett, J. L. (1972), 'The user interface in interactive systems', **Annual Review of Information Science and Technology**, 7, 159-196.

Bertrand-Gastaldy, S. and Davidson, C. H. (1986), 'Improved design of graphic displays in thesauri - through technology and ergonomics', **Journal of Documentation**, 42, 225-251.

Bivins-Noerr, K. T. and Noerr, P. L. (1982), The display device: a user oriented intelligent terminal, in: **Proceedings of the 6th International Online Information Meeting, London, 1982**, Learned Information, Oxford.

Boll, J. (1952), **The input and output of 22 psychological periodicals: a study of bibliographic coverage**, University of Illinois, Urbana, Illinois.

Borgman, C. L. (1984), 'Psychological research in human computer interaction', **Annual Review of Information Science and Technology**, 19, 33-64.

Bovey, J. D. and Brown, P. J. (1987), 'Interactive document display and its use in information retrieval', **Journal of Documentation**, 43, 125-377.

Brenner, L. P., Huston-Myamoto, M., Self, D. A., Self, P. C. and Smith, L. C. (1981), 'User computer interface design for information systems: a review', **Library Research**, 2, 63-73.

Brittain, J. M. (1970), **Information and its users: a review with special reference to the social sciences**, Oriel, Bath University Press, Bath.

Brittain, J. M. (1979), 'Information services and the structure of knowledge in the social sciences', **International Social Science Journal**, 31, 711-728.

Brooks, H. M. (1986), Developing and using problem descriptions, in: Brooks, B. C. ed. **Proceedings IRFIS 6 Conference, Frascati, Italy, 1985,** (Intelligent information systems for the information society), North Holland, Amsterdam.

Brooks, H. M., Daniels, P. J. and Belkin, N. J. (1985), Problem description and user models: developing an intelligent interface for document retrieval systems, in: **Proceedings of Informatics 8,** (Advances in intelligent retrieval), Aslib, London.

Carbonell, J. G. (1983), **The role of user modelling in natural language interface design,** Mellor University, Carnegie.

Celoria, F. (1968), 'The archaeology of serendip', **Library Association Record, 70,** 251-253.

Cheydleur, B. F. (1961), 'Information retrieval - 1966', **Datamation, 7,** 21-25.

Claparede, E. (1910), 'L'unificationet la fixation de la terminologie psychologique', **Archives de Psychologie, 9,** 109-124.

Cleverdon, C. W. (1960), **Report on the first stage of an investigation into the comparative efficiency of indexing systems,** College of Aeronautics, Cranfield.

Cleverdon, C. W. (1962), **Report on the testing and analysis of an investigation into the comparative efficiency of indexing systems,** College of Aeronautics, Cranfield.

Cleverdon, C. W. (1967), 'The Cranfield tests on index language devices', **Aslib Proceedings, 19,** 173-194.

Cleverdon, C. W. (1971), 'Design and evaluation of information systems', **Annual Review of Information Science and Technology, 6,** 41-73.

Cleverdon, C. W. (1972), 'On the inverse relationship of recall and precision', **Journal of Documentation,** 28 195-201.

Cleverdon, C. W., Mills, J., and Keen, E. M. (1966), **Factors determining the performance of indexing systems,** (Volume 1: Design), College of Aeronautics, Cranfield.

Cleverdon, C. W., and Keen, E. M. (1966), **Factors**

determining the performance of indexing systems, (Volume 2: Test Results), College of Aeronautics, Cranfield.

Croft, W. B. (1985a), 'An expert assistant for a document retrieval system', in: **RIAO 85 Actes of the Conference: Recherche d'informations assistee par ordinateur, Grenoble, France**, IMAG, Grenoble.

Croft, W. B. (1985b), **An expert assistant for a document retrieval system**, University of Massachusetts, Massachusetts.

Croft, W. B. (1986), User specified domain knowledge for document retrieval, in: Rabiitti, F. ed. (1986), **Proceedings of the 9th International Conference on Research and Development in Information Retrieval**, ACM, Washington.

Cronin, B. (1981), 'The need for a theory of citing', **Journal of Documentation**, 37, 16-24.

Cronin, B. (1982), 'Invisible colleges and information transfer: a review and commentary with particular reference to the social sciences', **Journal of Documentation**, 38, 212-236.

Cronin, B. (1984), **The citation process: the role and significance of citations in scientific communication**, Taylor Graham, London.

Cuadra, C. A., Katter, R. V., Holmes, E. H., and Wallace, E. M. (1967), **Experimental studies of relevance judgements: final report**, (Volume 1, Report no. TM-3520/001/00), Systems Development Corporation, Santa Monica; California.

Curtis, P. M. J., Pickering, H. and White, M. (1981), 'ALS update', **Vine**, 39, 34-36.

Dahlberg, I. (1978a), 'Progress in launching INTERCONCEPT', **International Classification**, 5, 102.

Dahlberg, I. (1978b), 'A referent-oriented analytical concept theory for INTERCONCEPT', **International Classification**, 5, 142-151.

Dahlberg, I. (1980), 'Classification and the social sciences', **International Classification**, 7, 55.

Dahlberg, I. (1981), 'Conceptual definitions for INTERCONCEPT', **International Classification**, 8, 16-22.

Daniels, P. J. (1986a), 'Cognitive models in information retrieval - an evaluative review', **Journal of Documentation**, 42, 272-304.

Daniels, P. J. (1986b), The user modelling function of an intelligent interface for document retrieval systems, in: Brooks, B. C. ed. **Proceedings IRFIS 6 Conference, Frascati, Italy, 1985, (Intelligent information systems for the information society)**, North Holland, Amsterdam.

Daniels, P. J., Brooks, H. M. and Belkin, N. J. (1985), Using problem structures for driving human-computer dialogues, in: **RIAO 85 Actes of the Conference: Recherche d'informations assistee par ordinateur, Grenoble, France, IMAG, Grenoble.**

Davies, C. C. (1978), 'Reference retrieval by user-negotiated term frequency ordering within a dynamically adjusted notional 'document'', **Journal of Informatics**, 2, 62-76.

De Mey, M. (1980), The relevance of the cognitive paradigm for information science, in: Harbo, O. and Kajberg, L. eds. **Theory and application of information research**, Mansell, London.

Diesing, P. (1971), **Patterns of discovery in the social sciences**, Aldine, Chicago.

Doyle, L. B. (1963), Is relevance an adequate criterion in retrieval system evaluation? in: **Proceedings of the American Documentation Institute 26th Annual Meeting**, 199-200.

Ellis, D. (1982), **Terminology of communication and terminology of retrieval in the social sciences**, MA Dissertation, University of Sheffield, Sheffield.

Ellis, D. (1984a), 'The effectiveness of information retrieval systems: the need for improved explanatory frameworks', **Social Science Information Studies**, 4, 261-272.

Ellis, D. (1984b), 'Theory and explanation in information retrieval research', **Journal of Information Science**, 8, 25-38.

Ellis, D. (1986), 'Social science information research', **Journal of the American Society for Information Science**, 37, 86-88.

Ellis, D., Roberts, N., Hounsell, D., Saracevic, T. and Persson, O. (1985), 'Comments on Wersig and Windel', **Social Science Information Studies**, 5, 25-32.

Ellis, D., Hendry, I. G., Nelis, K., Willett, P. and Wood, F. E. (1987), End-user searching of bibliographic databases: an evaluation of an experimental best match retrieval system', in: Dorrington, L. ed. **Online**

information retrieval in practice, Proceedings of the 2nd UK Online User Group Conference, Bristol, 1986. Taylor Graham, London.

Ford, G. (1977), **User Studies: an introductory guide and select bibliography**, (Centre for Research on User Studies (CRUS) occasional paper no. 1. BLRDD Report no. 5375), Centre for Research on User Studies, University of Sheffield, Sheffield.

Foskett, D. J. (1974), **Classification and indexing in the social sciences**, Butterworths, London.

Frei, H. P. and Jauslin, J. F. (1983), 'Graphical presentation of information and services: a user oriented interface', **Information Technology Research and Development**, 2, 23-42.

Garfield, E. (1963), 'Citation indexes in sociological and historical research', **American Documentation**, 14, 289-291.

Garfield, E. (1964), 'Citation indexing: a natural science literature retrieval system for the social sciences', **American Behavioural Scientist**, 7, 58-61.

Garvey, W. D. and Griffith, B. C. (1972), 'Communication and information processing within scientific disciplines: empirical findings for psychology', **Information Storage and Retrieval**, 8, 123-136.

Garvey, W. D., Lin, N. and Nelson, C. E. (1970), 'Communication in the physical and social sciences', **Science**, 170, 1166-1173.

Garvey, W. D., Lin, N. and Nelson, C. E. (1971), 'A comparison of scientific communication behaviour of social and physical scientists', **International Social Science Journal**, 23, 256-272.

Garvey, W. D., Lin, N., Nelson, C. E., and Tomita, K (1972), 'Research studies in patterns of scientific communication: I general description of research program', **Information Storage and retrieval**, 8, 111-122.

Garvey, W. D., Gottfred, S. D. and Simmons, J. G. (1984), 'A comparison of two major scientific information exchange processes in psychology: 1962-1976', **American Psychologist**, 39, 11-21.

Gerould, A. C. and Warman, H. J. (1954), 'Most cited periodicals in geography', **Professional Geographer**, 6, 6-12.

Glaser, B. (1978), **Theoretical sensitivity: advances in the methodology of grounded theory**, Sociology Press, Mill Valley, California.

Glaser, R. G. and Strauss, A. L. (1967), **The discovery of grounded theory: strategies for qualitative research**, Aldine, Chicago.

Green, T. R. G., Payne, S. J. and van der Veer, G. C. eds. (1983), **The psychology of computer use**, Academic Press, London.

Greene, R. J. (1977), 'The effectiveness of browsing', **College and Research Libraries**, 38, 313-316.

Griffiths, A., Luckhurst, H. C. and Willett, P. (1986), 'Using inter document similarity information in document retrieval systems', **Journal of the American Society for Information Science**, 37, 3-11.

Gull, C. D. (1956), 'Seven years of work on the organisation of materials in the special library', **American Documentation**, 7, 320-329.

Harter, S. P. (1971), 'The Cranfield 2 relevance assessments: a critical evaluation', **Library Quarterly**, 41, 229-243.

Heine, M. H. (1980), The 'question' as a fundamental variable in information science, in: Harbo, O. and Kajberg, L. eds. **Theory and application of information research**, Mansell, London.

Heine, M. D. (1981), Simulation and simulation experiments, in: Sparck-Jones, K. ed. (1981), **Information retrieval experiment**, 179-198, Buterworths, London.

Herner, S. (1970), Browsing, in: Kent, A. and Lancour, H. eds. **Encyclopedia of Library and Information Science**, (vol. 3.), Dekker, New York.

Hildreth, C. R. (1982), The concept and mechanics of browsing in an online library catalogue, in: Williams, M. E. and Hogan, J. H. eds. **National Online Meeting Proceedings 1982**, Learned Information, Medfor, New Jersey.

Hildreth, C. R. (1984), 'Online catalogues and public libraries', **Public Libraries**, 23, 59-60.

Hills, P. J., Hull, J. and Pullinger, D. (1983), **An**

experiment on the redesign of journal articles for online viewing, BLRDD, London.

Hjerppe, R. (1985), Project HyperCatalog: visions and preliminary conceptions of an extended and enhanced catalog, in: Brooks, B. C. ed. **Proceedings IRFIS 6 Conference, Frascati, Italy, 1985,** (Intelligent information systems for the information society), North Holland, Amsterdam.

Hogeweg-de Haart, H. P. (1983), 'Characteristics of social science information: a selective review of the literature: part I', **Social Science Information Studies, 4,** 147-164.

Hogeweg-de Haart, H. P. (1984), 'Characteristics of social science information: a selective review of the literature: part II', **Social Science Information Studies, 4,** 15-30.

Hollnagel, E. (1984), 'Experimental design of man machine studies', **Social Science Information Studies, 4,** 149-166.

Hollnagel, jE. and Woods, D. D. (1983), 'Cognitive systems engineering: new wine in new bottles', **International Journal of Man Machine Studies, 18,** 583-600.

Hounsell, D., Payne, P. and Willett, I. (1977), 'Searching ERIC: a small scale examination of the potential value of the ERIC database to educationalists in Britain', **Education Libraries Bulletin, 20,** 1-15.

Hutchins, W. J. (1978), 'The concept of 'aboutness' in subject indexing', **Aslib Proceedings, 30,** 172-181.

Hyman, R. J. (1972), **Access to library collections: an enquiry into the validity of the direct shelf approach with special reference to browsing,** Scarecrow, Metuchen.

Hyman, R. J. (1982), **Shelf classification research: past, present - future?,** University of Illinois, Illinois.

Ingwersen, P. (1982), 'Search procedures in the library analysed from the cognitive point of view', **Journal of Documentation, 38,** 165-191.

Inwersen, P. (1984), 'Psychological aspects of information retrieval', **Social Science Information Studies, 4,** 83-95.

Jakobovits, L. A. and Osgood, C. E. (1967), 'Connotations of twenty psychological journals to their

professional readers', **American Psychologist**, 22, 792-800.

Kaplan, A. (1964), **The conduct of Inquiry**, Chandler, San Francisco.

Katter, R. V. (1970), **On-line user of remote access citation retrieval services**, (Report TM(L)-4494/000/00), System Development Corporation, Santa Monica, California.

Kemp, D. A. (1974), 'Relevance, pertinence and information system development', **Information Storage and Retrieval**, 10, 37-47.

Kirsch, I. (1977), 'Psychology's first paradigm', **Journal of the History of the Behavioural Sciences**, 13, 317-325.

Kircz, J. G. and Bleeker, J. (1987), 'The use of relational databases for electronic and conventional publishing', **Journal of Information Science**, 13, 75-89.

Krantz, D. L. (1965), 'Research activity in 'normal' and 'anomalous' areas', **Journal of the History of the Behavioural Sciences**, 1, 39-42.

Krantz, I. (1972), 'Schools and systems: the mutual isolation of operant and non-operant psychology as a case study', **Journal of the History of the Behavioural Sciences**, 8, 86-102.

Lancaster, F. W. (1968), **Evaluation of the MEDLARS demand search service**, National Library of Medicine, Bethesda, Maryland.

Lancaster, F. W. (1979), **Information retrieval systems: characteristics, testing and evaluation**, 2nd. ed., Wiley, New York.

Lancaster, F. W. and Mills, J. (1964), 'Testing indexes and index language devices, **American Documentation**, 15, 4-13.

Lawrence, G. S. and Oja, A. K. (1980), **The use of general collections at the University of California: a study of unrecorded use, at-the-shelf discovery and immediacy of need for materials at the Davis and Santa Cruz campus libraries: final report**, Library Studies and Research Division, University of California, California.

Lesk, M. E. and Salton, G. (1968), 'Relevance assessments and retrieval system evaluation, **Information Storage and Retrieval**, 4, 343-359.

Levine, M. M. (1969), 'An essay on browsing', RQ, 9, 35-36, 93.

Lin, N. and Garvey, W. D. (1972), 'Information needs and uses', **Annual Review of Information Science and Technology**, 7, 5-37.

Lin, N. and Nelson, C. E. (1969), 'Bibliographic reference patterns in some sociological journals 1965-1966', **American Sociologist**, 4, 47-50.

Line, M. B. (1971), 'The information uses and needs of social scientists: an overview of INFROSS', **Aslib Proceedings**, 23, 412-434.

Line, M. B. (1976), Concluding summary, in: **International Workshop of Secondary Service Producers**, (BLRD Report 5289), British Library, Boston Spa.

Maass, S. (1983), Why systems transparency?, in: Green, T. R. G., Payne, S. J. and van der Veer, G. C. eds. (1983), **The psychology of computer use**, 19-28, Academic Press, London.

Maguire, M. (1982), 'An evaluation of published recommendations on the design of man-computer dialogues', **International Journal of Man Machine Studies**, 16, 236-261.

Marcus, R. S., Benefeld, A. R. and Kugel, P. (1971), The user interface for the INTREX retrieval system, in: Walker, D. ed. **Interactive bibliographic search: the user/computer interface**, AFIPS Press, Montvale, New Jersey.

Markey, K. and Atherton, P. (1978), **ONTAP: online training and practice manual for ERIC data base searchers**, ERIC Clearinghouse on Information Resources, New York.

Maron, M. E. and Kuhns, J. L. (1960), 'On relevance, probabilistic indexing and retrieval', **Journal of the Association for Computing Machinery**, 7, 216-244.

Martin, T. (1980), Information retrieval, in: Smith, H. T. and Green, T. R. G. eds. **Human interaction with computers**, Academic Press, London.

Martin, T. H. and Parker, E. B. (1971), Designing for user acceptance of an interactive bibliographic search facility, in: Walker, D. ed. **Interactive bibliographic search: the user/computer interface**, AFIPS Press, Montvale, New Jersey.

Martin, T. H., Carlisle, J. H. and Treu, S. (1973), '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, 142-147.

Martyn, J. (1974), 'Information needs and uses', **Annual Review of Information Science and Technology**, 9, 3-23.

Martyn, J. (1975), 'Citation analysis', **Journal of Documentation**, 31, 290-297.

Martyn, J. (1986), **Literature searching habits and attitudes of research scientists**, British Library, BLRDD, London.

McAllister, C. and Bell, J. M. (1971), 'Human factors in the design of an interactive library system', **Journal of the American Society for Information Science**, 22, 96-104.

Meadows, J. (1987), Communication in the scientific community, in: White, B. ed. **Information for all: access and availability**, Proceedings of the annual conference of the IIS, Peebles, 1986. Taylor Graham, London.

Momenee, K. (1982), 'A notation for describing the components of library information systems', **Journal of the American Society for Information Science**, 33, 270-280.

Moran, T. P. (1981), 'Guest editors introduction: an applied psychology of the user', **ACM Computing Surveys**, 13, 1-11.

Morehead, D. R. and Rouse, W. B. (1982), 'Models of human information seeking', **Information Processing and Management**, 18, 193-205.

Morehead, D. R. and Rouse, W. B. (1983), 'Human computer interaction in information seeking tasks', **Information Processing and Management**, 19, 243-253.

Morehead, D. R. and Rouse, W. B. (1985a), 'Online assessment of the value of information for searches of a bibliographic database', **Information Processing and Management**, 21, 83-101.

Morehead, D. R. and Rouse, W. B. (1985b), 'Computer-aided searching of bibliographic databases: online estimation of the value of information', Information Processing and Management, 21, 387-399.

Moser, A. (1978), 'Information concepts', **Journal of Documentation**, 34, 350-351.

Mote, L. J. B. (1963), 'Review of Cleverdon, C. W., The

Cranfield 1, 1962, report', **Journal of Documentation**, 19, 80-81.

Muravcsik, M. J. and Murugesan, P. (1975), 'Some results on the function and quality of citations', **Social Studies of Science**, 5, 86-92.

Nelis, K. (1985), **Human interaction with computers in an information retrieval context: a study of the users interaction with INSTRUCT**, MSc Dissertation, University of Sheffield, Sheffield.

Noerr, P. L. (1979), Information navigation, M. and Gray, K. eds. (1979), **Proceedings of Informatics 5**, (The analysis of meaning), Proceedings of a conference held by the Aslib Informatics Group and the BCS Information Retrieval Specialist Group), 202-205, Aslib, London.

Noerr, P. L. and Noerr, K. T. B. (1985), 'Browse and navigate: an advance in database access methods', **Information Processing and Management**, 21, 205-213.

Oddy, R. N. (1977a), 'Information retrieval through man-machine dialogue', **Journal of Documentation**, 33, 1-14.

Oddy, R. N. (1977b), 'Retrieving references by dialogue rather than by query formulation', **Journal of Informatics, 1, 37-53.**

Oddy, R. N. (1981), Laboratory tests: automatic systems, in: Sparck-Jones, K. ed. (1981), Information retrieval experiment, 156-178, Butterworths, London.

Paisley, W. J. (1968), 'Information needs and uses', **Annual Review of Information Science and Technology**, 3, 1-30.

Palay, A. J. and Fox, M. S. (1981), Browsing through databases, in: Oddy, R. N., Robertson, S. E., van Rijsbergen, C. J., and Williams, P. W., eds. **Information Retrieval Research**, Butterworth, London.

Patton, M. Q. (1980), **Qualitative evaluation methods**, Sage, London.

Pejtersen, A. M. (1984), 'Design of a computer-aided user-system dialogue based on an analysis of users search behaviour', **Social Science Information Studies**, 4, 167-183.

Pejtersen, A. M. and Austin, J. (1983), 'Fiction retrieval: experimental design and evaluation of a search system based on users' value criteria', **Journal of Documentation**, 39, 230-246.

Pejtersen, A. M. and Austin, J. (1984), 'Fiction retrieval: experimental design and evaluation of a search system based on users' value criteria', **Journal of Documentation**, 40, 25-35.

Pinsky, L. (1983), What kind of "dialogue" is it when working with a computer, in: Green, T. R. G., Payne, S. J. and van der Veer, G. C. eds. (1983), **The psychology of computer use**, 29-40, Academic Press, London.

Pratt, A. D. (1978), 'Information concepts', **Journal of Documentation**, 34, 242-244.

Prabha, C. G. (1983), 'Some aspects of citation behaviour: a pilot study in business administration', **Journal of the American Society for Information Science**, 34, 202-206.

Price, D. J. De Solla (1961), **Science since Babylon**, Yale University Press, Yale.

Price, D. J. De Solla (1966), 'Collaboration in an invisible college', **American Psychologist**, 21, 1011-1017.

Putnam, H. (1978), **Meaning and the moral sciences**, R. K. P., London.

Ramsey, H. R. and Grimes, J. D. (1983), 'Human factors in interactive computer dialogue', **Annual Review of Information Science and Technology**, 18, 29-59.

Rees, A. M. (1965), 'The Aslib-Cranfield test of the Western Reserve University indexing system for metallurgical literature: a review of the final report', **American Documentation**, 16, 73-76.

Rees, A. M. (1967), 'Evaluation of information systems and services', **Annual Review of Information Science and Technology**, 2, 63-86.

Rich, E. (1979), 'User modelling via stereotypes', **Cognitive Science**, 3, 329-354.

Rich, E. (1983), 'Users are individuals: individualising user models', **International Journal of Man-Machine Studies**, 18, 199-214.

Richmond, P. A. (1963), 'Review of the Cranfield project', **American Documentation**, 14, 307-311.

Riggs, F. W. (1979a), 'Importance of concepts - some considerations on how they might be designated less ambiguously', **American Sociologist**, 14, 172-185.

Riggs, F. W. (1979b), 'In search of rational

organization - introduction: shifting meanings of the term 'bureaucracy'', **Internation Social Science Journal**, 31, 563-584.

Riggs, F. W. (1979c), 'INTERCONCEPT aims redefined', **International Classification**, 6, 178-179.

Riggs, F. W. (1979d), 'A new paradigm for social science terminology', **International Classification**, 6, 150-158.

Roberts, S. A. (1981), 'Developing a focus for library and information research in social sciences in the United Kingdom', **Education Libraries Bulletin**, 24, 1-19.

Robertson, S. E. (1977), 'Progress in documentation: theories and models in information retrieval', **Journal of Documentation**, 33, 126-148.

Robertson, S. E. (1979), Between aboutness and meaning, in: MacCafferty, M. and Gray, K. eds. (1979), **Proceedings of Informatics 5**, (The analysis of meaning), Proceedings of a conference held by the Aslib Informatics Group and the BCS Information Retrieval Specialist Group), 202-205, Aslib, London.

Robertson, S. E. (1981), The methodology of information retrieval experiment, in: Sparck-Jones, K. ed. (1981), **Information retrieval experiment**, 9-31, Butterworths, London.

Robertson, S. E. and Belkin, N. J. (1978), 'Ranking in principle', **Journal of Documentation**, 34, 93-100.

Robertson, S. E. and Sparck-Jones, K. (1976), 'Relevance weighting of search terms', **Journal of the American Society for Information Science**, 27, 129-146.

Robertson, S. E., Thompson, C. L., Macaskill, M. J. and Bovey, J. D. (1986), 'Weighting, ranking, and relevance feedback in a front-end system', **Journal of Information Science**, 12, 71-75.

Ross, J. (1983), 'Observations of browsing behaviour in an academic library', **College and Research Libraries**, 44, 269-276.

Rouse, W. B. and Rouse, S. H. (1984), 'Human information seeking and design of information systems', **Information Processing and Management**, 20, 313-318.

Rouse, W. B., Rouse, S. H., and Morehead, D. R. (1982), 'Human information seeking: online searching of bibliographic citation networks', **Information Processing and Management**, 18, 141-149.

Rowland, J. F. B. (1982), 'The scientist's view of his information system', **Journal of Documentation**, 38, 38-42.

Ruckmich, C. A. (1913), 'The use of the term 'function' in English text-books of psychology', **American Journal of Psychology**, 24, 99-123.

Russett, B. M. (1970), 'Methodological an theoretical schools in international relations', **American Academy of Political Science Monographs**, 10, 87-105.

Salton, G. ed. (1971), **The SMART retrieval system: experiments in automatic document processing**, Prentice-Hall, Englewood Cliffs.

Salton, G. (1973), 'Recent studies in automatic text analysis and document retrieval', **Journal of the Association for Computing Machinery**, 20, 258-278.

Samelson, F. (1978), 'From 'race psychology' to 'studies in prejudice': some observationsa on the thematic reversal in social psychology', **Journal of the History of the Behavioural Sciences**, 14, 265-278.

Sanderson, N. (1984), **Graphics in information retrieval**, MSc Dissertation, University of Sheffield, Sheffield.

Saracevic, T. ed. (1970a), **Introduction to information science**, Bowker, New York.

Saracevic, T. (1970b), Ten years of relevance experimentation: a summary and synthesis of conclusions, in: **Proceedings of the American Society for Information Science 33rd Annual meeting**, 33-36.

Saracevic, T. (1975), 'Relevance: a review of and a framework for the thinking on the notion in information science', **Journal of the American Society for Information Science**, 26, 321-343.

Saracevic, t. (1978), Problems of question analysis in information retrieval, in: Brenner, E. H. ed. **Proceedings of the 41st ASIS Annual Meeting, New York**, (The information age in perspective], Knowledge Industry Publications, White Plains, New York.

Schuman, B. A. (1983), Interactive amiability: DIALOG, ORBIT, and BRS under scrutiny, in: Vondran, R. F. ed. **Proceedings of the 46th ASIS Annual Meeting, Washington, D. C.**, (Productivity in the information age), Knowledge Industry Publications, White Plains, New York.

Sharp, J. (1964), 'Review of the Cranfield-WRU test

literature', **Journal of Documentation**, 20, 170-174.

Sharp, J. R. (1967), 'Content analysis, specification and control', **Annual Review of Information Science and Technology**, 2, 87-122.

Shephard, R. N. (1962a) 'The analysis of proximities: multi-dimensional scaling with and unknown distance function', **Psychometrika**, 27, 125-140.

Shephard, R. N. (1962b) 'The analysis of proximities: multi-dimensional scaling with and unknown distance function', **Psychometrika**, 27, 219-246.

Skelton, B. (1971), **Comparison of results of science user studies with investigation into information requirements of the social sciences**, (DISISS working paper no 1), Bath University Library, Bath University, Bath.

Skelton, B. (1973), '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, 138-156.

Smelser, N. J. (1969), The optimum scope of sociology, in: Bierstedt, R. ed. **A design for sociology: scope, objectives, and methods**, American Academy of Political an Social Science, Philadelphia.

Smithson, S. (1987), **Experiments in the judgement of relevance**, (unpublished paper delivered at the BCS Information Retrieval Special Interest Group 9th Research Colloquium, Newcastle-upon-Tyne Polytechnic, 1987).

Sparck-Jones, K. (1979a), 'Search term relevance weighting given little relevance information', **Journal of Documentation**, 35, 30-48.

Sparck-Jones, B. (1979b), 'Experiments in relevance weighting of search terms', **Information Processing and Management**, 15, 133-144.

Sparck-Jones, K. (1981a), The Cranfield tests, in: Sparck-Jones, K. ed. (1981), **Information retrieval experiment**, Butterworths, London.

Sparck-Jones, K. (1981b), Retrieval system tests 1958-1978, in: Sparck-Jones, K. ed. (1981), **Information retrieval experiment**, Butterworths, London.

Sparck-Jones, K. and van Rijsbergen, C. J. (1976), 'Information retrieval test collections', **Journal of Documentation**, 32, 59-75.

Stenstrom, P. and McBride, R. B. (1979), 'Serial use by social science faculty: a survey', **College and Research Libraries, 40**, 426-431.

Stone, S. (1982), 'Humanities scholars information needs and uses', **Journal of Documentation**, 38, 292-313.

Summitt, R. K. (1971), 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**, AFIPS Press, Montvale, New Jersey.

Swanson, D. R. (1965), 'The evidence underlying the Cranfield results', **Library Quarterly**, 35, 1-20.

Swanson, D. R. (1971), 'Some unexplained aspects of the Cranfield tests of indexing language performance', **Library Quarterly**, 41, 223-228.

Swift, D. F., Winn, V. A. and Bramer, D. A. (1974), **A case study in indexing and classification in the sociology of education**, (Report for the period September 1970 - June 1973, 2 vols.), Open University, Milton Keynes.

Swift, D. F., Winn, V. A. and Bramer, D. A. (1977a), **A multi-modal approach to indexing and classification', International Classification**, 4, 90-94.

Swift, D. F., Winn, V. A. and Bramer, D. A. (1977b), 'Multi modality in indexing and searching', **Journal of Informatics**, 1, 91-95.

Swift, D. F., Winn, V. A. and Bramer, D. A. (1978a), ''Aboutness as a strategy for retrieval in the social sciences', **Aslib Proceedings**, 30, 182-187.

Swift, D. F., Winn, V. A. and Bramer, D. A. (1978b), **A sociologist's view of information systems**, (unpublished paper read to a meeting of the Insititute of Information Scientists, 1978).

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

Taylor, R. S. (1968), 'Question negotiation and information seeking in libraries', **College and Research Libraries**, 29, 178-194.

Thompson, D. A. (1971), 'Interface design for an interactive retrieval system: a literature survey and a research description', **Journal of the American Society for Information Science**, 22, 363-373.

Thorne, R. G. (1955), 'The efficiency of subject catalogues and the cost of information searches', **Journal of Documentation, 11,** 130-148.

Thouless, R. H. (1939), Problems of terminology in the social sciences, in: Bartlett, Sir F., Ginsberg, M., Lindgren, E. J., and Thouless, R. H. eds. **The study of society: methods and problems,** RKP, London.

Timbie, M. and Coombs, D. H. (1969), **An interactive information retrieval systems: case studies on the use of DIALOG to search the ERIC document file,** Stanford University, Stanford.

Tonnies, F. (1899a), 'Philosophical terminology', (Translated: Bosanquet, Mrs B.), **Mind, N.S., 8,** 289-232.

Tonnies, F. (1899b), 'Philosophical terminology', (Translated: Bosanquet, Mrs B.), **Mind, N.S., 8,** 467-491.

Tonnies, F. (1900), 'Philosophical terminology', (Translated: Bosanquet, Mrs B.), **Mind, N.S., 9,** 46-61.

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

Urquhart, D. J. (1976a), 'National lending/reference libraries or libraries of first resort', **BLL Review, 4,** 7-10.

Urquhart, D. J. (1976b), 'University libraries: the case for a national lending library system', **Times Higher Education Supplement,** 17th September, 8.

Van Rijsbergen, C. J. (1981), Retrieval effectiveness, in: Sparck Jones, K. ed. (1981), **Information Retrieval Experiment,** Butterworths, London.

Vickery, B. C. (1966), 'Review of Cleverdon, Mills and Keen (1966)', **Journal of Documentation,** 22, 247-249.

Vickery, B. C. (1967), 'Review of Cleverdon and Keen (1967)', **Journal of Documentation,** 23, 338-340.

Warren, H. C. (1911), 'Terminology', **Psychological Bulletin,** 8, 20-21.

Warren, H. C. (1912), 'Terminology', **Psychological Bulletin,** 9, 35.

Warren, H. C. (1913), 'Terminology', **Psychological Bulletin,** 10, 18-19.

Warren, H. C. (1916), 'Terminology', **Psychological Bulletin**, 13, 15-16.

Warren, H. C. (1918), 'Terminology', **Psychological Bulletin**, 15, 15-16.

Warren, H. C. (1920), 'Terminology', **Psychological Bulletin**, 17, 16-17.

Warren, H. C., Calkins, M. W., Dunlap, K., Gardiner, H. N. and Ruckmich, C. A. (1918), 'Definitions and delimitations of psychological terms I', **Psychological Bulletin**, 15, 89-95.

Warren, H. C., Calkins, M. W., Dunlap, K., Gardiner, H. N. and Ruckmich, C. A. (1922), 'Definitions and delimitations of psychological terms II', **Psychological Bulletin**, 19, 232-235.

Warren, H. C., Calkins, M. W., Dunlap, K., Gardiner, H. N. and Ruckmich, C. A. (1925), 'Definitions and delimitations of psychological terms III', **Psychological Bulletin**, 22, 370-374.

Watson, L. E., Gammage, P., Grayshon, M. C., Hockey, S., Jones, R. K. and Oldman, D. (1973), 'Sociology and information science', **Journal of Librarianship**, 5, 270-283.

Weimar, W. B. and Palermo, D. S. (1973), 'Paradigms and normal science in psychology', **Science Studies**, 3, 211-244.

Weinstock, M. (1971), Citation indexes, in: Kent, A. and Lancour, H. eds. **Encyclopedia of Library and Information Science**, (vol. 5.), Dekker, New York.

Wersig, G. (1970), 'Terminologies and classification in the social sciences', **Nachrichten fur Dokumentation**, 21, 259-262.

Wersig, G. (1971), **Information-Kommunikation-Documentation**, Verlag Dokumentation, Munchen.

Wersig, G. (1979), The problematic situation as a basic concept of information science in the framework of the social sciences - a reply to N. Belkin, in: **New trends in informatics and its terminology**, Viniti, Moscow.

Wersig, G. and Windel, G. (1985), 'Information science needs a theory of information actions', **Social Science Information Studies**, 5, 11-23.

Wilkins, L. T. (1964), **Social Deviance**, Tavistock Publications.

Wilson, P. (1973), 'Situational relevance', **Information Storage and Retrieval**, 9, 457-471.

Wilson, T. D. (1980), 'On information science and the social sciences', **Social Science Information Studies**, 1, 5-12.

Wilson, T. D. (1981), 'On user studies and information needs', **Journal of Documentation**, 37, 3-15.

Winn, V. A. (1971), 'A case study in the problem of information processing in the socia science field: the OSTI-SEA project', **Aslib Proceedings**, 23, 76-88.

Wittgenstein, L. (1967), **Zettel**, Oxford, Blackwell.

Wormell, I. (1984), 'Cognitive aspects in natural language and free text searching', **Social Science Information Studies**, 4, 131-141.

Wright, K. (1973), **Social science information characteristics with particular reference to the Educational Resources Information Centre (ERIC)', Journal of the American Society for Information Science**, 24, 193-204.

Xhignesse, L. V. and Osgood, C. E. (1967), 'Bibliographical citation characteristics of the psychological journal network in 1950 and 1960', **American Psychologist**, 22, 778-791.

Appendix One

Research/Teaching Interests of the Social Scientists Interviewed

MRC/SSRC Social and Applied Psychology Unit

Research Interests

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

Research/Teaching Interests

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 and Continuing Education

Research/Teaching Interests

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 pre-school; 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, Economic and Social History, Geography, Politics, Sociology, Prehistory and Archaeology

Research/Teaching Interests

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 Two

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 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?

Do you follow up references cited in material consulted?

How do you decide which references to follow and when to stop?

How do you decide which references to cite in your own work?

What is the most difficult problem you experience in looking for material or keeping up-to-date?

General

Have you ever made use of any indexing or abstracting service?

Have you found them useful?

Have you ever used a citation index (Science Citation Index/ Social Science Citation Index)?

Did you find it useful?

Have you ever made use of the Current Contents?

Did you find them useful?

Have you ever had an online (computer based) search carried out?

Did you find the results useful?