USER REDEFINITION OF SEARCH GOALS THROUGH INTERACTION WITH AN INFORMATION RETRIEVAL SYSTEM

PHILIP MARTIN HIDER

for the degree of

DOCTOR OF PHILOSOPHY

submitted to

CITY UNIVERSITY, LONDON

DEPARTMENT OF INFORMATION SCIENCE
SCHOOL OF INFORMATICS

JULY 2004

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The author would like to thank his principal supervisor, Prof Stephen Robertson, for his patience and invaluable advice, and who has been a continual companion on this long journey, despite the physical distance that has separated us.

The author is also indebted to OCLC, Inc., in providing transaction logs from its FirstSearch service, on which the main analysis described in this thesis is based, and in particular, Mr Andrew Wang and his staff in the division of OCLC Asia Pacific.

Grateful thanks are conveyed to Mrs Ho-Lee Kiat Pia, for permitting the author to conduct the pilot study in Tampines Regional Library, Singapore.

Special mention must be made of Ms Koo Lik Yuen and Ms Lim Wan Hua, who assisted the author with the pilot study in a most effective manner.

Grateful thanks are also conveyed to various service managers who corroborated certain factual information concerning key FirstSearch databases.

The author is particularly appreciative of the encouragement and consideration he has received from his employers and former employers during the pursuit of the work represented in this thesis.

The thesis has been completed only because of the unrelenting support the author has received from his family and friends, and most of all, from his wife and parents.

DECLARATION

I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to me. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

2/12/04

Philip Martin Hider Date

ABSTRACT

Search goals of users of information retrieval systems have commonly been assumed to be static and well-formed. However, a significant amount of goal redefinition is detected in the studies described. A pilot study examined user behaviour at a library OPAC, showing that search results would quite frequently induce users to reconsider and revise their search goals, sometimes following up with a new search based on this revision (labeled "strong" goal redefinition). The main analysis employed transaction logs from the OCLC FirstSearch service, investigating what factors, if any, might affect the amount of goal redefinition that takes place during a search session. To this end, ten hypotheses were proposed and considered. Within each search session, logged queries were coded according to their conceptual differences or similarities, in order for indices of strong goal redefinition to be constructed: a chronological content analysis was thus performed on the transaction logs. The indices of redefinition for search sessions on different FirstSearch databases were compared. It was found that different databases induced goal redefinition to different extents. Further analysis showed that the metadata displayed by a database appeared to affect the amount of goal redefinition, and that the presence of abstracts in results was a positive factor, as was the presence of descriptors and identifiers, perhaps because of the former's hyperlinking nature on the FirstSearch interface. On the other hand, no evidence was found to indicate that abstract length has much of an effect on goal redefinition, nor hit rate or levels of precision and recall. Of the two indices of redefinition that were produced, the "refined" index showed signs of greater precision. Implications of the findings are discussed. It is suggested that goal redefinition should be considered a positive result of system feedback, and that systems should readily allow users to follow up on redefined goals. Abstracts and summaries of documents should be presented to the user as frequently as possible, and hyperlinks from key terms in the metadata should also be created to assist evolving searches. The importance of how system feedback is encountered by the user is emphasized in a new model of information retrieval, which embraces the nonconscious as part of the "cognitive viewpoint," allowing for nonconscious information wants to enter into a user's consciousness through cues encountered during the scanning of search results, triggering a redefinition of search goal. This thesis paves the way for a considerable amount of potentially important research, including: further testing and development of the index of goal redefinition; deeper transaction log analyses, perhaps using screen recorders, examining semantic content and contextualizing at the level of the query; and further identification and analysis of the factors affecting goal redefinition, across different types of information retrieval system.

Chapter 1 INTRODUCTION

This study initially centred around the theme of "bibliographic serendipity," a term coined by the author to represent what he saw as an underestimated phenomenon in information retrieval. Originally, bibliographic serendipity was employed to mean the *accidental* discovery of bibliographic data that might lead the user to valuable information. However, it was soon reasoned that a *purely* accidental discovery of bibliographic data cannot, by definition, be improved upon (assuming that such improvement would be a "good thing") through system design or user training. As soon as the system designer or user *aims* for more "serendipitous" discovery, then we move away from pure chance. Even when the user is offered the genuinely random retrieval of documents or their record surrogates (a feature available on some Web search engines), the system designer is assuming *some* information intention.

Instead, a weaker notion of bibliographic serendipity was developed, in which metadata encountered by the user does not represent exactly what the user had previously conceived of as their information goal, but which is nevertheless related to it. As such, the encounter is not a pure accident; yet it triggers a review of what the user's information wants actually are, and perhaps lead to a revision of their current search goal. This revision is referred to in this thesis as "goal redefinition."

Goal redefinition can occur through interaction with all kinds of information retrieval systems, although this study focuses on systems which retrieve on metadata, or bibliographic information (in the first instance). Search goals are redefined *incidentally,* rather than accidentally, and when this leads to the acquisition of information, what we have, the author contends, is a form of incidental information acquisition.

Such information acquisition forms an important component of information seeking, but has rarely been studied in information retrieval, probably because of the emphasis on purposeful searching and fixed information goals and relevance criteria used to evaluate system performance. The author wished to investigate

whether in fact interaction with an information retrieval system might not also include important elements of incidental information acquisition. Users do not simply search; they also browse or scan results, and the system feedback that they encounter may reshape their information wants and goals, as well as meet or deny them.

It has been commented how a browsing facility on an information system can help facilitate "serendipity." Yet browsing takes place not only when such a facility is used; it can also occur when results produced by *matching* are presented to the user. Might not this also result in serendipity?

A fundamental premise of goal redefinition through interaction with a system is that the user does not necessarily start a search session with a perfect conceptualization of their information goal, that there is room for revision. A reference interview might sharpen a search goal prior to a session, and so might feedback from the system during the course of a session. While information goals and relevance judgements might be fixed in laboratory tests, in real life, users' goals may evolve as they become conscious of aspects that they were not previously conscious of. In this way, the author is of the view that the nonconscious and the changing state of a user's cognition need to be brought into a more dynamic model of user-system interaction.

While it may be clear, to the author at least, that goals are sometimes redefined through interaction with information retrieval systems, little research has investigated the extent to which this happens. If it happens very infrequently, then there is less of a need to account for changing states of cognition and to examine which factors might affect goal redefinition. Thus the first aim of this study is to discover whether, at least in some information retrieval contexts, the phenomenon's prevalence is significant. A pilot study indicated that this was indeed the case, with OPAC (online public access catalogue) users in a library apparently redefining their search goal on perhaps one third of occasions. A distinction was made following the pilot study between "strong" and "weak" redefinition: the former was where a revised goal was followed up on with another search (on the same system).

The pilot study was based on interviews conducted before and after search sessions, and on recordings of the sessions using the screen logger, ScreenCam. However, it was found that these methods of investigation were not very scaleable, given the fragility of the software and the shyness of the library users. A different source of data was sought for the main analysis, which attempted to analyse the factors that might be involved in search goal redefinition. The transaction logs from the OCLC FirstSearch service were made available to the author for this purpose.

Large numbers of search sessions were represented on the FirstSearch logs, even after considerable filtering of non-applicable sessions. The search sessions used in the analysis were on different FirstSearch databases with different content characteristics. The search interface, on the other hand, was uniform across the databases, so comparison of the amount of goal redefinition produced by sessions on the different databases could be made. Two indices of goal redefinition were constructed in order to do this, based on the coding of each query in the sessions. This coding, performed by the author (following confirmation of its reliability by sample parallel coding), might be described as a form of content analysis, taking account of the context of the search session. Primarily based on semantic elements, rather than syntactic ones, the coding interpreted at a deeper level than most previous transaction log analyses have done.

The indices of goal redefinition produced results that supported their reliability, with similarities between search sessions on the same databases, yet differences between search sessions on different databases. However, an important assumption was made: the amount of strategic reformulation of queries was reasonably uniform across databases, at least with respect to certain types of query.

The results based on the indices were subject to various statistical tests. Differences were clearly detected across certain kinds of databases. First, the item-level database, WorldCat, induced goal redefinition much less than did the other databases, which represented article-level documents. Second, full-text

databases induced considerably more redefinition than did those which did not offer full text. Third, databases with abstracts appeared to encourage redefinition, compared with those without abstracts. Fourth, descriptors seemed to facilitate redefinition, probably due, at least in part, to their hyperlinking nature.

On the other hand, length of abstract was not found to particularly affect the amount of redefinition, nor levels of precision or recall, or hit rate, or the subject represented by a database. However, this does not mean that these variables have no bearing on goal redefinition, and further research is needed, across different systems, to establish more general conclusions about the weight different system and user characteristics have on goal redefinition.

Chapter 2 RESEARCH QUESTION

2.1 Incidental Information Acquisition and Search Goal Redefinition

As was mentioned in the previous chapter, the subject of this study was originally phrased as "bibliographic serendipity," and more explicitly, the accidental acquisition of information during the document retrieval process. The stumbling block was the concept of *accidental.* If information acquisition is a *complete* accident, then it must be independent of any document retrieval system; yet what the author wished to study, in particular, were ways in which a system might *affect* "serendipity." It was thus realized that a modified definition of serendipity was necessary: the *incidental* acquisition of information *through* the document retrieval process.

Information *incidentally* acquired does not satisfy the initial search goal of the system user, but nevertheless pertains to an information need or want. It could be that the information relates to a completely different information want, or it could be that in actual fact it relates to the initial search goal. In the latter case, what has happened is that the initial search goal has been redefined before, or upon, encountering this information, such that the information's utility is seen in the light of a new consciousness. It is this new consciousness which makes the process "serendipitous."

The incidental acquisition of information that nevertheless relates to the initial search goal is a kind of serendipity that may well be affected by aspects of an information retrieval system. A key element of this phenomenon was search goal redefinition, and although this phenomenon would not necessarily lead to information acquisition, since a redefined goal may not be stated, it was considered a worthy subject of study in its own right. Search goal redefinition has been little researched, and may well be influenced by system design. Moreover, goal redefinition was assumed to be a positive aspect of information retrieval, something that benefited information seekers. *Search goal redefinition* thus became the subject of this study.

2.2 Search Goals, Information Goals, Tasks and Problems

We use the terms *search goal* and *information goal* in this study to represent the same thing: *the aim of the user when interacting with an IR system.* While the immediate search goal in a document retrieval system is to find the document which contains the information, the ultimate aim is to find the information itself. If users redefine their search goal, they are in effect redefining their information goal – search goal is dependent on information goal.

"Goal" is used in this study as a neutral concept, and could be based on an information "task" or an information "problem." Thus the goal may be to solve the problem or to fulfill the task; in both cases this is done by finding the appropriate information. As Beaulieu (2000) points out, investigators of information seeking have tended to develop models at a "task-level," whereas information retrieval research has tended to focus on lower levels of abstraction, but the author does not consider it necessary to limit an investigation in this way – tasks and problems are constructs imposed on what is essentially a continuum.

2.3 Information and Document Retrieval Systems

While serendipity and information goal redefinition pertains to information seeking in general, the author is particularly interested in its occurrence during interaction on a particular kind of information system, namely, document retrieval systems. Such a system is defined here as one which retrieves records that represent other documents (which may or may not be immediately available to the user of the system), and in this study is also called a *bibliographic* retrieval system. Information goal redefinition probably occurs through all types of information retrieval system, but this study focuses on bibliographic systems, those with which the author is most familiar.

2.4 Browsing and Searching

As we note in chapter 3, the literature commonly identifies browsing as an information seeking behaviour that fosters "serendipity." What this generally means is that a non-specific search goal – such as "something interesting" – is met when a specific information want is recognized ("a book on birds' eggs is just what I'm looking for"). This recognition might then lead to a search *goal* if further information relating to the specific information want is desired. In this way, browsers may quite often redefine their information goal from the very broad to the much narrower.

One way therefore in which document retrieval systems could facilitate goal redefinition would be to provide browsing functionality. However, this might not be the only way. Browsers are able to identify specific information wants by evaluating the information they encounter. Yet users submitting matching-type queries also evaluate the information they encounter. For example, they evaluate items (or their surrogates) in result sets. Such evaluation may be carried out according to a fixed conceptualization of an information goal, but it does not need to be, and in real life the author supposes that it frequently is not. Whether the frequency of redefinition during *matching-type searching* is affected by aspects of an information system, is one of the key questions to be addressed in this study. It is a question which has rarely been discussed before.

2.5 Goal Redefinition Questions

The three primary questions that needed to be answered were thus:-

(a) does goal redefinition produced by interaction with document retrieval systems, occur in *significant* amounts? (We shall examine what we mean by significant later.)

(b) does such redefinition occur when searching on some systems more than it does when searching on others?

(c) could the *amount* of redefinition be measured in a way (or ways) that allows for investigation of the first two questions?

The third question begs two more: what do we mean here by (i) amount and (ii) measure?

2.6 Quantification of Goal Redefinition

By the amount of redefinition, we could mean the extent to which the redefinition moves away from the original definition, or we could mean the number of redefinitions occurring (or we could mean both). While an index representing an objective measure of the difference between one goal and another was considered, particularly in terms of subject goals (perhaps based on a classification scheme or the author's ranking), it became apparent that such an index would be seriously unreliable, given the subjective nature of subject classification and the immensely varying contexts of users' information circumstances and conceptual relationships between data in document retrieval systems. This leaves the *frequency of redefinition* to represent – we assume accurately – the amount of redefinition.

2.7 Redefinition as a Construct

Herein lies another question: is redefinition a discrete variable? More attention needs to be given to the nature of redefinition. In real life, the information seeker often does not articulate redefinition – yet it still occurs. We know it occurs because we experience it ourselves, and recall such experience. In this way, redefinition is a *construct* of the searcher. As a construct, it may be either discrete or continuous. The researcher could ask the user to indicate each *time* their goal has been redefined, which would represent a construct discrete in nature.

2.8 Measurement of Redefinition

This leads us to the other definitional question, measurement. The user may articulate redefinition, as a discrete variable, for the benefit of the researcher, but will such articulation be an *accurate* representation of the user's construction? We should note that any expression of redefinition, whether consciously or unconsciously produced by the user, may exclude some construction which the seeker failed to express, for whatever reason. The point is that we must be satisfied that any method of measurement we adopt produces a reasonable representation of the frequency of redefinition, or at least provides an *index* of redefinition that portrays, with reasonable accuracy, *relative* frequencies of redefinition.

As the index of search redefinition was being devised, it was realised that redefinition of information goals involving *non-topical* aspects would be particularly difficult to identify. Indeed, non-topical redefinition was not accounted for in the refined index (see section 5.13) at all. We therefore need to assume that a focus on topical redefinition does not distort the indices.

2.9 Three Basic Hypotheses

The three primary questions for investigation listed above in section 2.5 were reformulated as the following hypotheses.

(1) Information seekers redefine their information goals as a result of interaction with some bibliographic information retrieval systems significantly frequently.

(2) An index can be produced that accurately represents the relative frequencies at which information seekers redefine their information goals as a result of interaction with various bibliographic information retrieval systems.

(3) Information seekers redefine their information goals as a result of interaction with some bibliographic information retrieval systems more frequently than they do with others.

Hypothesis 2 has been placed before 3, since the investigation of 3 is necessarily dependent on the acceptance of 2. Similarly, if hypothesis 1 is rejected, then hypotheses 2 and 3 become immaterial.

2.10 Additional Hypotheses on Possible Redefinition Factors

If hypothesis 3 is accepted, then the next question, of course, is *why* do some systems encourage more redefinition than do others? We have already noted that systems which allow for browsing may well encourage more redefinition compared with those which do not. Other possible factors include (i) retrieval mechanism, (ii) interface, and (iii) displayed content. All three of these areas are worthy of study; the author's selection of additional research questions was based on methodological considerations rather than on any predilection towards particular system factors.

Chapter 5 discusses methodological issues in detail; suffice to state here that the nature of the research subject necessitated a methodology based on a real-life information-seeking situation (as opposed to a simulated one), and when using real-life bibliographic systems, controlling displayed content is especially problematic, and opportunities to study the effect of a range of retrieval mechanisms or interfaces are rare. For pragmatic reasons, therefore, the study of possible system factors on redefinition focused on differences in displayed content.

When the opportunity of using data derived from the use of a particular set of databases presented itself – namely the OCLC FirstSearch databases – the differences in displayed content chose themselves: the author identified what appeared to be the major differences in displayed content between the databases, and the following four hypotheses were formulated, based on these differences.

(4) Differences in the nature of the displayed content of bibliographic information retrieval systems affect the frequency at which information seekers redefine their information goals as a result of interaction with the systems.

(5) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems which display abstracts more frequently than they do with bibliographic information retrieval systems which do not display abstracts (all other aspects being equal).

(6) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems which display longer abstracts more frequently than they do with bibliographic information retrieval systems which display shorter abstracts (all other aspects being equal).

(7) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems which display descriptors more frequently than they do with bibliographic information retrieval systems which do not display descriptors (all other aspects being equal).

We note that hypothesis 4 is the base hypothesis here – hypotheses 5-7 cannot be accepted if hypothesis 4 is rejected.

Three other possible factors presented themselves for study, given the data the author decided to use for the main analysis. The first might be termed "search failure," though more accurately it concerns *hit rate.* It was speculated that low numbers of hits may encourage users to reconsider their goal, and thus to redefine them (when appropriate). This would only be partly a system factor; it would also be a use factor (failure may be the result of misuse as much as poor system performance). Given that hypotheses 5-7 are based on the theory that users encountering bibliographic content (they deem worth reading) are more

likely to redefine their goals (that is, sharpen their focus), to the extent that low numbers of hits generally provide less bibliographic content, the above speculation is the antithesis of hypotheses 5-7. If these three hypotheses were rejected, therefore, it might be because hit rate is a larger factor than content, in which case we would expect the following hypothesis to be accepted.

(8) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems more frequently when such interaction reveals fewer relevant documents (in the view of the seeker).

Another variable turned into a hypothesis for study was "retrievability." The hypothesis below is based on retrieval mechanisms available to users of the FirstSearch database.

(9) Levels of precision and recall produced by a bibliographic information retrieval system affect the frequency at which its users redefine their information goals.

Finally, another hypothesis was formulated relating to another aspect of content, and also use, namely discipline.

(10) Differences in the disciplines covered by different bibliographic information systems affect the frequency with which their users redefine their information goals.

2.11 Meaning of Significance

Another word used in hypothesis 1, "significantly," still needs to be discussed. Search goal redefinition may certainly occur, but does it do so to a *significant* extent? Similarly, a factor may affect redefinition, but nowhere near as much as other factors – is the factor still *significant?* The question, then, is what do we mean by significance? We can apply statistical tests of significance, and will do so, but these do not imply significant prevalence; rather, they allow us to be

satisfied that there *is* a phenomenon or a factor. For example, when testing for differences in frequency of redefinition, a statistical test resulting in a p value of 0.0005 would convince us that a particular factor were a real one. But this does not mean that it is necessarily a "significant" one, even though at one level, any factor is worth noting because it *could* be addressed if the opportunity presents itself. If there were no costs associated with addressing the factor, such that redefinition increased, then the factor would be significant, however small (assuming that redefinition is a positive aspect of user-system interaction, a matter further discussed in chapter 8).

The problem, of course, is that there are invariably *some* costs associated with every aspect of an IR system. If the benefits accrued are very limited, then it may not be worth the cost of implementing, using and maintaining a particular feature. However, it is almost impossible to quantify these potential costs. How can one tell what the costs of an implementation might be in the future, given the unknown technological and economic changes that will occur? A finding might not be significant for today's industry, but it might well be for tomorrow's. In fact, it is hardly any easier to quantify potential benefits, particularly in relation to different systems.

For hypotheses 2-10, therefore, we are concerned with *existence* rather than significance. Whether the discovery of any factors might have a serious impact on contemporary system design is discussed in chapter 8. Hypothesis 1 includes the word "significantly" because the author wishes not only to confirm existence of the phenomenon of search redefinition itself – this is clear to many experienced online searchers – but to consider its prevalence, a matter which appears to have been overlooked. The decision as to the benchmark we should use here is obviously a subjective one, but the author regards it as reasonable to assume that some system designers would be interested in the phenomenon if over ten per cent of search goals were revised through interaction with a typical bibliographic system.

2.12 Restriction to Subject Searching

Matching-type searching, involving the submission of *queries,* was the kind of information seeking on which this study focused. Indeed, for methodological reasons, it was decided to limit the main analysis further, to *subject searching,* which shall be defined in detail in chapter 5. While even known-item searches might result in redefinition, this was considered by the author more difficult to detect from the data available to him, and that the indices were indirect – and thus weak – enough already.

Chapter 3 LITERATURE REVIEW

3.1 Models of Information Retrieval Behaviour

Information retrieval has often been criticised for its systems-based approach, but until recently the message was not always getting across. Instead, another strand of information science developed which focused on information seeking behaviour, often outside the confines of specific information retrieval systems. The approach taken by the information seeking researchers has only in the past decade begun to be employed to study use of particular IR systems.

Earlier critiques of "hard" IR models failed to produce a paradigm shift. For example, Brookes (1981) argues that the traditional measures in IR may be too simplistic, based on an objectivity unsuited to the essentially subjective nature of information. He points out that relevance is dependant on one's knowledge base and advocates that the linear scales should be replaced with logarithmic scales and more sophisticated measures. More complex mathematical formulae have subsequently been introduced, but these do not necessarily account for the user any better than the simpler ratios.

Borgman (1996) spoke for many when she complained of users' frustration and ignorance at the OPAC terminal. While "hard" IR was beginning to see its efforts translated into more sophisticated commercial systems, with the Internet as a primary driver, the link between research and actual user benefits was still not always convincing. Borgman (1996:493) saw the problem as that of a lack of user perspective in traditional IR research, arguing that "online catalogs continue to be difficult to use because their design does not incorporate sufficient understanding of searching behaviour." Saracevic (1996) says much the same thing, while Yee (1991) and others had been calling for more user-friendly OPACs since the early 1990s, if not before.

Borgman (1996:494) was fighting against what she saw as the dominance of the "matching" model in IR system design: "most online catalogs are based on query design that allow some degree of search modification but are far from being

exploratory systems." In the traditional model, a system is evaluated from an "objective" viewpoint, with goals and relevance predetermined and static. Systems attempt to match input and output; but they do not attempt to *develop* input conceptually. The "matching" objective may suffice for laboratory systems, but not necessarily in real life.

Borgman argues that the traditional IR model can be traced back even before the early information science of the 1950s and 1960s, as part of a longer tradition of assumptions about the catalogue user. Cutter's famous objectives of the catalogue (1904), for instance, "made assumptions about the knowledge people brought to the information-seeking process" (Borgman 1996:495). The focus on system was perhaps understandable given the more limited flexibility that earlier technology allowed. However, the sophistication of the programming that system designers now have at their disposal, has inevitably advanced the cause of user-oriented design. IR systems have become much more flexible in recent years, and more interactive. It is probably true, as Saracevic contends, that IR systems have in fact overtaken theory with respect to interactivity. "Interaction practice is flourishing, interaction research is not" (Saracevic 1996:8). Certainly, designers have developed systems with a user-friendliness unimaginable to previous generations of catalogue user. The question is whether the researchers and their IR models are catching up. New models of information retrieval with a more user-oriented bent have been proposed, and it is those that will be discussed next.

User-oriented IR models

Bates' "berry-picking" model (1989) was one of the first to challenge the traditional matching model. The model has the user, and not the system, as its focus. The user has a range of strategies at his disposal, of which entering a query into an IR system is only one. The model emphasizes the idea that the user need not complete his information seeking in one throw; instead, he may pick off different elements of a task bit by bit. In this way, the information goal itself may evolve. Bates (1989:409) observes that the iterative IR systems were still very much based on static information needs and situations:

> The presumption is that the information need leading to the query is the same, unchanged, throughout, no matter what the user might learn from the documents in the preliminary retrieved set... The point of feedback is to improve the representation of a static need, not to provide information that enables a change in the need itself.

Unfortunately Bates makes little attempt to elaborate on how the *underlying* information need may change in her berry-picking model. While she paints a picture of an evolving search, produced through interaction with various information resources, this evolution seems more the result of changes of strategy than of *overall* goal. The whole information seeking process is broken down into segments whereby different subgoals are met through different search strategies; that the overall search problem may be reconceptualised by the seeker, is not brought out in Bates' model.

Ellis (1989) also felt that system designers had previously failed to question whether (and how) a system constrains the information seeker's behaviour, although a few experimental systems had been more accommodating (for example, the I^3R system produced by Croft & Thompson 1987). Ellis (1989) advocated a behavioural approach to IR system evaluation, investigating first the behaviours exhibited during the seeking of information, and then examining whether systems allow and promote these types of behaviour. Six types were identified: starting, chaining, browsing, differentiating, monitoring, and extracting.

Ingwersen (1992) advocated a "mediator model," in which interaction between system and user was key, and recommended that systems be more flexible and accommodate more varied forms of information seeking behaviour, which in Ingwersen's terms, meant offering both "open" and "fixed" subsystems, the former showing context and allowing for successful browsing, the latter allowing for matching and accommodating "retrieval."

A similar model to that of Bates' berry-picking one, upon which a theory was developed, describes "information foraging," as defined by Cronin and Hert

(1995). This "ecological" model aims to redress the balance between "matching" techniques and those less exacting techniques such as browsing. Cronin and Hert point out that "the best match principle" whereby "outcomes are matched against pre-set expectations" may not be applicable in many real-life situations. Instead, they adopted Sandstrom's concept of "scholarly foraging" (2001), similar to Bates' berry-picking idea. Information foragers are situated in a "hunter-gatherer" environment: some information seeking is carried out through matching (hunting); other times, foraging (gathering) techniques are used.

An information foraging theory is set out in more detail by Pirolli and Card (1999). The idea was that information seekers behave such that they optimise their behaviour in a way that gives them the best possible information "diet." This is a functionalist approach in which information "nutrients" are sought in "patches" on a cost-benefit basis. Information foragers would select a patch according to the density of the nutrients available, and the effort required to "travel" to it. Once a patch had become sufficiently depleted and another patch sufficiently attractive (in comparison with staying where they were), they would move on.

Pirolli and Card's theory (1999) contextualises information seeking, so that the attractiveness of an information resource changes over time, as other resources come into view. We may be able to enhance the "scent" – represented, for example, by bibliographic citations – that an IR system gives out to a user, or indeed we can enhance the user's power of "smell." However, the theory does not really incorporate changing goals, that is, a forager's diet is not said to change; as well as dietary revision due to the discovery of new types of "food," it could be that the forager only realises at a later stage exactly what type of food the patch could yield.

Cronin and Hert (1995) considered the theory particularly applicable to the way users interact with the Internet, and less manageable document space than those used in traditional IR experiments. Browsing or "surfing" may often be a more suitable activity in such environments. In such cases, users are not certain what they can find, but they may not need to be. Sandstrom (2001) puts forward the

attribute of *novelty* as an important information "currency" in scholarly information.

If the traditional IR systems cater for hunters, with their resultant "hits," it is clear that we should also look at how best to support foraging activity, which has a different functionality. However, we should also remember that hunters can also be foragers and vice-versa; indeed information hunter-gatherers may switch behaviour at a moment's notice, if conditions suit.

The costs cited in the information foraging theory include relocation (from one patch to another) and energy expended finding nutrients. Specific *cognitive* costs are not identified. Harter (1992), on other hand, in his discussion of the psychological nature of relevance, notes that the "processing" costs of information are significant in many situations. A simple example: one may opt to skip a technical exposition of a problem if it looks too complicated (it's not "worth" knowing).

Saracevic (1996) considers the interactive models of Ingwersen (1992) and Belkin (1980) as a step in the right direction, but that they lack testability. He thus puts forward another model which presents different "levels" of interaction: surface, cognitive and situational. System and user interact at each of these levels, and each of these levels has a bearing on the other two. Saracevic (1997) refines this model to incorporate "query "and "affective" levels, and extends it to the key concepts of relevance and feedback, and, most interestingly from this author's point of view, to search term selection. What Saracevic (1997:323) appears to attempt here is to broaden the concept of the search process as a subject of study, so that changes at the surface level of query are interrelated to changes that occur at the cognitive and "higher" levels:

> The selection process is realized and manifested on the surface level, while the effectiveness of search terms, involving user relevance judgements, is established at the cognitive and possibly also the situational levels, with the affective level playing a significant role as well.

Sutcliffe and Ennis (1998) proposed another model from an HCI perspective.

Their synthetic theory is based on four main activities of the IR process: problem identification, need articulation, query formulation, and results evaluation. A familiar feedback cycle is developed (compare, for example, Spink 1997), whereby results evaluation may lead to query reformulation, but how the process should evolve if a problem is re-identified does not appear to be covered.

Another stratified model is presented by Mizzaro (1998). In his model of dynamic interaction, any of four levels may change: real information need, perceived information need, request, and query. The request is the articulation of the perceived need, the query is its articulation according to the rules of the system. Changes at any level may have a bearing on the other levels. Relevance, which is multifaceted, changes as information needs, both real and perceived, change. The model centres on the user's cognition, but it is not clear how perception of need is to be distinguished by the researcher from its articulation, nor how real and perceived needs interface.

Like Ingwersen and Saracevic, Beaulieu (2000) stresses interaction as the central subject of investigation for IR. Although she distinguishes between the models of information seeking, and those of information retrieval, or "information searching," Beaulieu (2000:437) points out that "potentially, all aspects of information seeking could [now] be undertaken in an online environment, thus breaking down the boundaries between information seeking, searching and retrieval research."

We shall now turn to some specific aspects of information seeking that are of particular relevance to the subject of this study.

3.2 Cognition

Cognition has been an important aspect of information seeking research for many years, although not often tackled head on. Dervin's sense-making model (1977) gives cognition primacy; her hermeneutic approach has spawned a constructivist strand of information seeking (see, for instance, some of the papers in Vakkari, Savolainen & Dervin 1997). Belkin's ASK (Anomalous State of Knowledge)

concept (1980) also emphasises the importance of the user's cognition, but takes a more functionalist line.

Cremmins' survey (1992:28) of the meaning given to "cognition" across disciplines indicates that it is used for "all *conscious* mental processes, acts, behaviors, faculties, functions, modes, powers, or states which allow people or individuals to become aware of and know and interpret their environments and objects" (author's italics). Harmon and Ballestaros (1997) diverge from this definition, demonstrating the importance of the unconscious element of "cognition" during information seeking. They are, however, in a very small minority amongst information scientists, the focus having been very much on conscious thought processes. The exclusion of the unconscious is seen by the author as a major limitation on our understanding of an information seeker's mental state. Conscious cognition represents only the tip of the iceberg of a person's knowledge. As I begin to write this sentence, I know that elephants have trunks, but am not *conscious* of this fact (the fact entered my consciousness mid-sentence). The mind stores a vast reservoir of *potential* cognition, and this reservoir is what we will later term the *nonconscious.*

3.3 Information Acquisition without "Seeking"

Incidental information acquisition has been shown to play a major part in everyday information "seeking." Erdelez (1997) has preferred to use a less purposeful term to describe such activity, namely "information encountering." She proposes four categories of encounter – super encounterers, encounterers, occasional encounterers, and non-encounterers – to indicate the different levels of susceptibility that people have to such information acquisition. She notes that information encountering need not be left purely to chance. "Super encounterers believe that they are able to create situations conducive to information encountering" (Erdelez 1997:417).

However, there have been few other studies of the phenomenon. Williamson (1998) studied incidental information acquisition amongst a group of elderly subjects in Australia, with particular attention given to acquisition through

telecommunications. Williamson uses what she calls an "ecological model" and examines the actors' overall information environment. She found that her subjects picked up a lot of useful information (to them) incidentally. When people interact with the world of information, they are clearly able to identify information that is useful even when they are not specifically, or purposefully, seeking it.

Erdelez (1997) calls for a fresh approach to IR systems design, one which fully acknowledges other kinds of information acquisition apart from the purposeful, matching search. She argues that it is "important that information systems developers become aware of information encountering and other methods of users' real life information acquisition when making decision on information presentation and accessibility" (Erdelez 1997:420). She supports the development and application of Bates' berrypicking model, which challenges the traditional IR model.

It is worth noting that Erdelez (1997:419) associates information encountering with browsing, suggesting that systems could encourage the former by making "the information contents 'more browsable.'" She also points out that systems should be designed not to inhibit information encountering, for example, through "technical limitations of Internet resources, especially low response time, that makes it impossible for them to effectively explore, and retrieve information" (Erdelez 1997:419).

Incidental information acquisition is represented by behaviours, such as browsing and informal communication, which provide more scope for the acquisition of new knowledge that has been less well-defined by the information seeker, where new connections and analogies are made between topics, and more creativity is generated (Bawden 1986).

3.4 Browsing and Selecting

A common form of information seeking behaviour that has been identified as an important alternative to the "matching" kind is browsing. Indeed, Borgman

(1996) has claimed that 30-45% of library document retrieval is carried out through browsing (on the shelves as well as in systems), although Nicholas (1996) does not find users browsing on a specialised online retrieval system to the same extent. Nicholas observed that even when end-users were given the option of browsing, they mostly performed short query-based searches. He concludes that end-users need more guidance to optimise their information seeking.

Bawden (1993) considered that the dominance of the matching model had sidelined browsing in both systems design and IR research. Since then, however, many systems have developed some sort of browse function, and the ARIST literature review on browsing carried out by Chang and Rice (1993) testifies to the increasing weight ascribed to this activity in research.

Bawden (1986, 1993) emphasises the importance of browsing for creativity and serendipity. He subdivides the activity into three kinds: "capricious" browsing, "purposive" browsing, and "semi-purposive" browsing. It is the last category which he considers the most common form, which would seem to equate to Herner's "semi-directed" browsing (1970). In this kind of browsing, there is some goal in mind, but not as specific as the one associated with "directed" browsing. In other words, it fits in the middle of a continuum, at the extremes of which are random searching and comprehensive scanning. Bawden (1993) argues that browsing is never truly random, with no goal in mind whatsoever. At the very least, the user is looking for something "interesting." On the other hand, a browse is usually not the same as a *search,* with a particular item of information sought.

As Hancock-Beaulieu (1990) points out, it is not that browsing is a less systematic cognitive process: research shows that the human eye browses and identifies relevant items in a methodical way. As such, browsing is dependent on displays of readily scannable data. The data may be pictorially represented: visual representation of documents and information has become a popular, though difficult, area of research since Bawden's article (1993), under the label of "visualisation."

In the modern online environment of graphical user interfaces, browsing and selecting are not restricted to one-dimensional lists; they are often much closer to the three-dimensional reality of browsing through the power of hypertext. Interfaces have come a long way since BROWSE, the prototype system developed by Paley and Fox (1981), which provided hyper-links to other "relevant" documents according to a range of user criteria.

Borgman (1996) points out that databases behind IR systems represent parts of a bibliographic universe that is essentially abstract – it cannot be represented in the same direct way as tins on a supermarket shelf. It can, instead, be represented metaphorically – as a bookshelf, as a map, and so on. Which metaphors can best represent this universe, or the contents of a specific database, is not clear, nor is the advantage that a system employing a particular metaphor has over a simple display of bibliographic citations. Despite the interest shown by researchers in visualisation of literatures and collections (see, for example, White & McCain 1997), information science remains ill-equipped to compare the effectiveness of different visual representations.

Nevertheless, some interesting work has shown that this area is very much worth pursuing. Examples of experimental visualizations include those developed in the VIBE and GUIDO systems (Olsen, Korfhage & Sochats 1993 and Korfhage 1991), which help users to visualise the relative merits of retrieved documents through the relative positions of their iconic surrogates on the screen; the map displays in Lin's studies (1995, 1997); the PACE graphical interface described by Beheshti, Large and Bialek (1996), which simulated images of books and library shelves; and the Book House system (Pejtersen 1993) which also did this, and other aspects of the physical library.

Studies such as these indicate that the way in which metadata is presented has *some* effect on selection and browsing behaviour. Clearly, there is a very significant psychological aspect to this question. Eisenberg and Barry (1988) show how even the order of presentation on the screen affects relevance judgement. If the user is thinking about subjects and subject relationships, a more

appropriate metaphor may relate to the "work" rather than to the physical manifestation (compare the current enthusiasm in the bibliographic community for the catalogue model proposed by FRBR (IFLA 1998), a report into the functional requirements of bibliographic records which distinguished between representations of items, manifestations, expressions and works). Ingwersen (1992) argues that what is needed is not one single solution to document representation, but systems capable of *polyrepresentation,* of offering optimal interfaces according to users' varied cognitive situations, their particular problem spaces. An early attempt to produce a more flexible, interactive system in this vein is described by Belkin, Marchetti and Cool (1993).

More conceptual aspects of feedback have also been found to affect browsing: large versus small categories; subject versus disciplinary arrangements; commonalities of authorship, subject headings, publisher, and so on. Browsing has been shown to be enhanced by hierarchical index systems with cross-referencing in alphabetical lists (Allen 1993).

The way in which the human eye scans citations, bookshelves, or other forms of display is still being investigated. Toms' study (1997) of navigation through a digital newspaper by means of various tools, indicates that browsing is assisted by a "stable orientating device" or some form of structure (perhaps based on a classification scheme); furthermore, the content itself can make a difference, with the browsers identifying "landmarks or cues" in the text. Toms pursues the "cues" concept further in her 2002 paper: cues can shift a user's search focus (Toms 2002:858). It seems possible that elements of bibliographic records could also act as cues.

Bawden (1993) observes that browsing need not be a separate activity, but often occurs in tandem with query-based searching, as when users browse their hit lists or leave the catalogues to browse the shelves. If browsing is what users do when they select citations and documents via a query, then we need to examine this activity more closely when considering the impact of feedback from a system on the user.

3.5 Serendipitous Discovery

We have previously mentioned (in chapters 1 and 2) that pure serendipity is a limited topic of study, but that incidental, as opposed to accidental, information acquisition may be a rich area for information behaviour research. A more practical use of the term "serendipity," then, might equate to incidental information acquisition, as discussed above. In this broader sense, serendipity has been mentioned in earlier papers, on occasion.

Liestman (1992) proposes six categories of serendipity ("timely but accidental discovery") that can happen when patrons use libraries. "Coincidence" is the purest kind – statistical chance. "Prevenient grace" is less accidental, where the patron or librarian has increased the odds through the prior organisation of information (e.g. subject classification). "Synchronicity" is a result of "hidden patterns and unknown forces aiding the researcher." "Perseverance" is due to seekers' determination to find information. "Altamirage" assumes that unique behaviors can cause serendipity. "Sagacity" is a result of a more pragmatic and applied approach on the part of the information seeker. What Liestman appears to be saying is that successful information seekers often make their own luck, and that we should denounce the stigma often attached to "serendipity" by academics.

Kranich et al. (1986) found that serendipity was reduced when catalogues first switched from card to online, though subject searching increased. Doll (1980:55) also warned of a "narrowing of serendipity" due to the increasing popularity of database research. Richardson (1981:44) argued that the problem was exacerbated by the need for an intermediary with many of the earlier systems:

> The requestor is denied the opportunity to refine or restructure the initial query as system responses modify the perception of the problem. It is probable that more good solutions have been discovered by accident than by intent, and in the case of an online, interactive retrieval system the loss of this serendipity factor can be especially significant.

Since the 1980s, attempts and suggestions have been made to build serendipity-inducing elements into IR systems. Correspondingly, evaluations of IR systems

have sometimes incorporated "serendipity" or a related concept as a criterion. One positive sign was that users began to show more interest in subject searching than they had in the days of the card catalogue (Matthews 1986). There was considerably more subject *access,* although users struggled to optimise this. Indeed, Larson (1991) and others have detected a shift towards general keyword searching (when such a facility became widely available).

As computing advanced, authors such as Rice (1988) expressed increasing confidence that the serendipity produced by browsing card catalogues could be reproduced in the online environment. However, it took a while longer before it actually was. In 1990, Hancock-Beaulieu was still calling for online catalogues that could support both matching and "contextual" approaches to searching. In her before-and-after study (1990), she noted that the library's manual PRECIS index supported contextual broad search formulations whereas the OPAC encouraged matching and narrow formulations. Although the "success rate" of the online catalogue was slightly better than that of the manual tools, fewer items were retrieved at the shelves, and non-users of the bibliographic tools appeared to be just as successful.

The advent of the Internet has in many ways overtaken this concern. By 1996, Haglund had identified the great serendipitous potential of cyberspace. "Surfing" the World Wide Web can produce many serendipitous discoveries. As Internet browsing becomes a commonplace activity for increasing numbers of people, so research interest in its potential as a medium for incidental information acquisition, or serendipity, is surely going to gain momentum.

Finally, we should note how browsing and "serendipity" have also been found to be positively correlated (e.g. Snelson 1994), just as browsing and information encountering have.

3.6 Relevance

The concept of relevance has always been key to IR research. However, the early studies on precision and recall tended to assume a fixed, "objective" definition of

relevance, and a dichotomy: a document was either relevant or not relevant, and it could not become more or less relevant. Such assumptions were convenient for laboratory testing, but they did not necessarily reflect the real world. In his longitudinal study, Smithson (1994) describes how the vast majority of citations initially thought applicable or "relevant" to a sample of research students turned out to be otherwise in the final analysis. Thus Smithson calls into question traditional evaluation measures of IR systems.

As Harter (1992:603) points out, relevance, in its everyday sense, is much more complex than simple topicality. Searchers judge documents in terms of how "interesting" they are, and this goes beyond their aboutness. There are clearly many criteria for whether a document is relevant, in the sense of *useful* (the sense originally recommended as a basis for a measure of *utility* by Cooper 1973). Pao (1989), amongst others, points out that not everything that is *topical* is necessarily "relevant." If it is *old* information, that with which the user is already familiar, then it is unlikely to benefit them, at least not directly – they cannot utilize it. Similarly, topical documents may be unreadable (in a foreign language, for example).

Since the 1990s, several researchers have examined the concept from the user's perspective. From her longitudinal study, Wang (1997) concludes that document retrieval systems need to help the user make more accurate relevance judgements by offering more than the obvious topical pointers. Wang found that when users handle the document itself, they examine tables of contents, ascertain author's expertise, and consider citation status.

Barry (1994) has explored the way in which relevance judgements are made according to various aspects of documents, as well as "situational" factors. In a further study, she found that the metadata elements provided clues to document relevance in the following order of importance: abstracts, titles, bibliographic citations, indexing terms (Barry 1998). The bibliographic elements provided clues to "relevance" in many ways.

In earlier research, Saracevic (1969) showed that relevance judgements were more accurate when based on citations with abstracts compared with citations without. The importance of abstracts as a basis for relevance judgements is also found in a study by Marcus, Kugel and Benenfeld (1978). Janes (1991:643) concurs, "clearly, the abstract is the most important and most used single piece of information in relevance judging." Marcus, Kugel and Benenfeld (1978) also found that longer abstracts were generally more useful than shorter abstracts. The content and quantity of the bibliographic record as a factor in the changing of information need, as well as in relevance judgement, will be analysed as part of this thesis.

It has been observed that the application of relevance judgements, in real life, can also depend on scanning fatigue. Wiberley, Daugherty and Danowski (1995) show how catalogue searchers are, on average, persistent for the first 30 to 35 postings (after which they give up, perhaps trying another search). However, the study took only the most persistent browse per search session – it failed to look at the dependency of searches within a session.

Spink and Greisdorf have been at the forefront of research into relevance judgement. In one study (Spink & Greisdorf 1997), they examined "partial relevance" and found that searchers were more likely to encounter documents they deemed partially relevant if they were to change their information problem definition; or if their interaction with the search system had improved their knowledge of the topic in question; or if they were to revise their criteria of relevance. Spink and Greisdorf (1997:323) concluded that "a relationship exists between partially relevant items retrieved and changes in the users' information problem or question during an information seeking process." Thus "partially relevant" documents might also be useful during the early stages of the process, and that the more vague the information problem, the more items were judged partially relevant and the more chance that the user's problem definition was revised. This conclusion ties in with the point made by Borgman (1996) and others, namely, that users often fail to fully conceptualise their goal in the initial stages.

Spink and Greisdorf were later joined by Bateman in a study (Spink, Greisdorf & Bateman 1998) that analysed users' relevance judgements in terms of both criteria and extent. "Partially relevant" judgements were correlated with particular relevance criteria. Differences between users' criteria for highly, partially and non-relevant items were also identified. This in turn pointed to certain bibliographic elements becoming more or less important as clues to relevance across different stages of the information seeking process, a phenomenon observed by Bruce (1994). Spink, Greisdorf and Bateman (1998:599) proposed "a useful concept of relevance as a relationship and an effect on the movement of a user through the iterative stages of their information seeking process."

A further study was described by Spink and Greisdorf (2001). They concluded that "utility" was a more important criterion than topicality for determining high relevance. They also observed how users' calibration of relevance differed when judgements using continuous and categorical scales were compared. Spink and Greisdorf questioned the soundness of the traditional precision ratio as a measure of relevance when many "partially relevant" and "partially not relevant" documents could be included or excluded, and when these values were interpreted differently by different users. Spink and Greisdorf (2001) propose another measure of retrieval effectiveness, the "median effect," which takes the non-dichotomous nature of real-life relevance judgements into account by using a standard measure of central tendency of the distribution of judgements made from relevant through partially relevant to irrelevant.

The dynamic nature of real-life relevance judgements is now quite firmly established in the literature. Schamber (1994) distinguishes between the "system view" of relevance, which assumes that a document has a fixed value of relevance; the "information view" of relevance, whereby relevance depends on a particular judger; and the "situation view" of relevance (or "pertinence") that not only depends on judger, but on the situation of the judger when they make a decision about relevance. In the situation view, judgements are affected by feedback from the system.

Approaches which recognise how definitions of relevance change are listed by Schamber (1994) and include Kuhlthau's model of the information search process (1991). Sutton (1994) also presents a model of changing relevance judgements as information seekers' "mental models" change. Tang and Solomon (1998) investigate an end-user's relevance judgements made at both citation and full-text stages of an information retrieval process. They show that not only do judgements change in relation to particular documents, but that the criteria for judgements change (Tang & Solomon 2001). Given that there are so many ways in which a document might be judged relevant – Schamber (1994) provides a sample list of no fewer than eighty – it is understandable that new criteria enter users' definitions as search sessions progress.

Nevertheless, a considerable amount of recent research does not take account of relevance redefinition, focusing on instead other aspects of relevance judgements. For example, while Barry (1998) underlines how relevance should be related to information goal as a whole rather than individual queries, and Greisdorf and Spink (2000) point out how in real life users think in terms of degrees of relevance, rather than the traditional relevant/not relevant dichotomy, in neither case is there discussion of the possibility that the information goal itself might evolve. Greisdorf and Spink (2000:390) point out, in fact, that degrees of relevance are not the same as ambiguity – it is not that users do not know what they want, they are just not 100% or 0% confident that their information want will be met by a particular document.

Recent investigations into relevance criteria, such as those by Tang, Shaw and Vevea (1999), Wang and Soergel (1998), and Barry and Schamber (1998), also fail to allow for relevance *change,* particularly at the information search level. New user-based measures of relevance, such as those developed by Greisdorf and Spink (2001), and Su (1998), are similarly static in this respect.

Finally, we should bear in mind that the importance of relevance *feedback* has been recognised by designers of many modern search engines, in which a "related items" function is often offered (although Spink, Jansen and Ozmultu (2000) found that it was not commonly utilised by Excite users). Just *how*

important user feedback is, remains an open question. Hancock-Beaulieu, Fieldhouse and Do (1995) developed a "game-board" interface for their Okapi system, whereby users were able to select their own terms for query expansion, yet this interface option did not prove to be especially popular, and its retrieval performance was deemed less effective, compared with previous tests featuring automatic query expansion. We should note, on the other hand, that the study assumed a static, specific search goal.

3.7 Information Needs and Wants

Directly related to relevance, is information need or, as Vickery (1997) would have it, and as the author would prefer, "information want." A relevance judgement is, in effect, a reflection of the user's information want, usually conceived in terms of a search goal. Before we examine the literature's view of search goals, however, we need to clarify the higher level concept of information want. We use the term "goal" in this study because the redefinition that we are to examine is based on a specific intention, whereas many information wants (probably the vast majority) are not acted upon at any given moment. Since all goals are assumed to be derived from information wants, we are interested in the literature's views of both goals and wants.

Few IR researchers have discussed the dynamic nature of information wants or needs. Harter (1992) is an important exception. In his paper on the psychological nature of information retrieval, he argues how information needs are dependent on context, which changes as new information becomes available and old information re-enters consciousness as it takes on new meaning in the light of new information.

> The theory suggests that since there is no absolute 'information need,' there will be no single goal motivating a search. Since the searcher's cognitive state changes and evolves with the discovery of each relevant citation, one might expect a correspondingly dynamic search process. As relevant citations are discovered, new Boolean combinations of existing concepts will be formed. Terms representing the ideas of interest may disappear, or new terms emerge. The inquirer may feel the need to construct wholly new concepts and search facets, representing fresh ideas or cognitive connections. Concepts, authors, disciplines, institutional affiliations, journals, publications dates, and other parts of the bibliographic record may be seen in novel ways as a search progresses, resulting in a radical revision of the search problem, and moving the search process in new directions, Or, based on the relevance of the output, the inquirer may become more or less confident

that the current approach taken is fruitful and will act accordingly (Harter 1992:611).

We shall be exploring exactly this phenomenon, in terms of the way in which content encountered as feedback from an IR system affects perception of one's search goal, which would thus change relevance judgement.

3.8 Information Goals

A critical point about information or search goals is that they that may not be perfect: users may be less than perfectly clear as to exactly what it is they are looking for. In fact, the distinction between information want, and the expression or awareness of this want, has long been recognised in information science. Yet until the advent of the expert system, the issue tended to be left for those involved in user education to deal with. Over the past two decades, some expert systems have been created to help users formulate their queries as accurately as possible. An early example was I^3R (Croft & Thompson 1987), which asked the user a series of questions before connecting to the bibliographic database; such a system replaces the benefits previously afforded by a human expert intermediary.

Such systems have rarely found their way into the average library, however. Hancock-Beaulieu (1993) found that the real-life user preferred browsing subject headings to performing subject keyword queries; the subject headings assisted in reformulation much more than did the rest of the system. Hancock-Beaulieu (1993: 278) points out that

> To improve search results, more attention needs to be given to the quality if interaction in the search process... The initial search strategy is but one stage of the search process. The reformulation of searches is proving to be more problematic.

Borgman (1996: 496) argues forcefully that the users' "conceptual knowledge" of how to translate an information need into a searchable query involves, at least in part, an evolving consciousness of the information need itself:

> [the process] often begins with some vaguely-felt need of wanting to know something and gradually evolves to the point where one can describe some attribute of the documents that might contain the information. Once the need can be phrased sufficiently to begin searching, the question itself may change through multiple interactions of finding and using information resources. Thus people usually approach an information

> retrieval system with a partially-formed query to be negotiated.

Borgman (1996:493) also notes that

> research in information seeking indicates that users formulate questions in stages, gradually coming to the point where they can begin to articulate a query... Yet the design of most operational online catalogs assumes that users formulate a fixed goal for the search and that each search session is independent.

Vague goals are not only found amongst more informal information seeking, when browsing shelves and the like. They are found amongst many users of sophisticated IR systems. Ingwersen (1996) may draw a matrix with well-defined and stable information needs producing "matching" information seeking, in one corner, and ill-defined, unstable information needs producing browsing, and so forth, in the other corner, but in real life, we cannot draw a clear-cut distinction between different information contexts. Toms (2002:857) observes that "users do not always have explicit goals such as 'look for information about ...' or 'learn something about ...' ... [A goal] may range from a precise, well-defined purpose to an imprecise, or ill-defined purpose." Hjørland (1997:160) declares that "well-functioning libraries and well-functioning information services not only satisfy needs that have already been articulated, but are also a part of the process of making it possible to acknowledge and articulate these needs in the first place." The same can be said for well-functioning IR systems.

Whatever, the degree of clarity, there has been a tendency in IR to treat search goals synchronically, even by researchers who have emphasised the interactivity of the IR process. A search goal is thus often a given on which a model is to be built or a system measured. For example, Tague-Sutcliffe's measure of "informativeness" (1992) is based on real-life interaction between user and system, but is nevertheless based on a static search goal. In a more recent example, Toms' model of "information interaction" (2002:857) combines user, content and system, but the information goal is a given, despite the fact that it "may range from a precise, well-defined purpose to an imprecise, or ill-defined purpose." Interestingly, Toms adds that users are driven to seek information "not necessarily because of a conscious admission of a void" (Toms 2002:858).

What has sometimes happened is that the information want has been stratified, allowing *subgoals* to be seen as changing and "situated," but where a higher level of goal (or problem) still defines the model, as a given – see, for example, Hert (1995, 1997) and Ng (2002). Likewise, Xie (2000) emphasises the different levels of user goal and how each may shift, but concentrates on the lowest level, "interactive intention." Xie (1997) shows how strategies may change during interaction with the system and that these changes may stem from changes in "interactive intentions." Her analysis identified seven types of interactive intention and several types of interactive intention shift. However, these shifts, such as "assisted shift" and "alternative shift," are themselves mostly strategic in nature. It is worth noting that Xie (1997:108) adds, "I also noticed shifts of current search goal, especially those shifts related to 'keeping up to date.' Shifts of current search goals need to be further analyzed."

In some cases, researchers have studied query reformulation *specifically* in terms of strategic adjustments, not adjustments or refinements of goal: when the user reformulates a query because the query's syntax is problematic, not because the query's underlying goal requires fine-tuning. Most IR experimentation involving "interactive systems" focuses on producing better matches between a set goal and texts. A typical example might be Debili, Fluhr and Radasoa (1989). Many transaction log analyses have studied search reformulation, but only at the level of strategy (see section 5.1.4 for a review). Shenouda (1990) also examined search "modifications," yet none of the categories of reasons he constructed was specifically related to search goal.

Nevertheless, the dynamic nature of information and search goals is beginning to be acknowledged by researchers, often when studying relevance. Wang and White (1995), for example, in their longitudinal study indicate how goals can change over time while describing how relevance judgements change as new information is encountered and criteria change. Vickery (1997) discusses some key concepts used in information science, including information want and need, query, relevance and information search, his presupposition no. 33 being: "If the enquirer is not yet ready to state 'I have been informed,' then query may be modified, either to express the original information want more clearly, or to

express a changed understanding of the nature of the want."

Query refinement has been linked to goal change by a few researchers, such as Wang and White, and Xie, *en passant,* although a lack of clarity remains: the distinction between query reformulation for strategic reasons and query reformulation because of a review of information want, is still rarely made. As Robertson and Hancock-Beaulieu (1992) ask, to what extent might query reformulations be due to new ideas through system interaction, and to what extent are expressions of an information want kept back until needed? This is a question which remains largely unanswered. Three studies discussed below, however, have provided clues that help to answer this question. They represent very different approaches, but are considered by the author to be the most significant research in the subject area of this thesis.

Shifts of focus

Robins (1997, 2000) examines shifts of focus on an information problem during dialogue between end-user and search intermediary, including broader/narrower/related term shifts (as Aigrain and Longueville (1994:225) observe in relation to possible iterative IR system techniques, queries may be modified either vertically, i.e. broader/narrower, or horizontally, i.e. laterally). Robins finds many shifts taking place, particularly during online interaction. These shifts include strategic shifts, but some shifts appear to be related to information goals, including those Robins (1997:128) codes as "SCOPE" ("determine the boundaries and scope of the desired information") or "TOPIC" (2000:920). These only account for 3%, however, while many more are "operational" in nature, instigated by the intermediary in order "to get enough information to structure the query without regard for the conceptual basis for the search" (Robins 1997:129). For example, those coded "STRAT" (term changes for strategic reasons) cover 24% (or 31% of shifts whilst online, according to Robins' 2000 paper). Robins (2000) goes on to examine the results of a survey carried out at the time of the dialogue and finds "moderate evidence" for changes of information problem.

It is very possible that the high level of strategy-related shifts would be lower if there were no intermediary. It is also worth noting that the information needs represented by the searches were characterised by Robins (1997:132) as "well-defined and stable." Robins (2000:926) thus points out that situation factors may play an important role in determining the extent of information problem change, and notes, that Kuhlthau's subjects (1991), for example, "were much younger and perhaps less skilled at formulating research questions." Robins (1997:132) concludes that "one of the remaining questions in information behaviour research is whether users' conceptions of their information problem changes during information retrieval interaction. Further research should, without question, attempt to identify such changes."

Reformulation and remembering

Another study that coded reformulations is described by Dalrymple (1990). Instead of dialogue, Dalrymple's coding is based on the search terms themselves, according to a protocol analysis (monologue). She compares initial statements of information need with those used in the search session. Dalrymple would also have used transaction logs had they been available.

Dalrymple's study is based on a theory of remembering from psychology, a cognitive process called "reformulation," and investigates how search and presearch reformulations may vary according to other variables, such as the nature of the IR system. In particular, she compares the amount (frequency) of reformulation produced by the card catalogue versus the online catalogue, for the same collection. The transcripts were coded in order to construct a count of reformulations per search (session). Dalrymple found that the card catalogue produced significantly fewer reformulations than did the OPAC (although it is questionable whether she should have employed a t-test to this end, given the non-Gaussian nature of the distribution and the very low means).

Dalrymple thus demonstrates that the nature of a system can have an effect on the amount of reformulation that takes place. However, Dalrymple (1990:274) does not elaborate greatly on the role reformulation may play in the IR process.

She defines it "as an interactive refinement of the term or terms used to locate bibliographic items in a catalog." According to the psychological model, "reformulations are prompted by examples or instances that are present in the searcher's environment, thus enabling the searcher to criticize or refine *what is being sought"* (author's italics). This would imply that reformulations indicate changes or refinement of search goal. But Dalrymple goes on to state that "the instances that are available in the searcher's environment enhance the searcher's understanding of the meaning of the *terms* used and remind searchers of likely *terms* to used in the retrieval process" (author's italics). This suggests that the feedback produced by the IR system also helps at the strategic level, providing new terms for the searcher to use, and it may be that while Dalrymple considers systems giving rise to more reformulations to be better "retrieval" systems, the main reason why she thinks this is so, is because of the strategic benefit. Unfortunately, Dalrymple does not distinguish between different types of reformulation – most crucially, between reformulation induced by strategic considerations and reformulation caused by a genuine shift of information want.

Dalrymple does, however, put forward a couple of factors to account for varying amounts of reformulation. First, she suggests that "the relationship between frustration and reformulation could be examined to determine whether reformulation is a response to frustration" (1990:278). Discussing her findings, Dalrymple (1990:278) considers "it unclear whether feedback from the external environment precedes reformulation, or whether the reformulation acts as a request for feedback from the system." Of course, there could well be an element of both taking place.

Second, Dalrymple (1990:278) notes how the *speed,* as well as the nature, of the feedback from the system may also have a bearing on the amount of reformulation that occurs. This may be compared with Erdelez's suggestion that information encountering may be encouraged when users are able to browse through content more swiftly (Erdelez 1997).

Unconscious cognition

Harmon and Ballestaros (1997) published a paper which appears quite unique in the information science literature: "Unconscious cognition: the elicitation of deeply embedded information needs." They point out that the role of "unconscious cognition" has been largely ignored in IR research (though it may be noted that Harmon himself has sometimes taken a rather "structuralist" approach to information retrieval, see e.g. Harmon 1970). Although many researchers might consider it a contradiction in terms, "unconscious cognition" is defined by Harmon and Ballestaros (1997:422) as "that set of cognitive contents and processes in the cognitive system that is unavailable to awareness but nevertheless affects thought and action."

While intelligent and expert systems have made some attempt to elicit the unconscious part of users' information needs, they have not so far been particular successful, and most research in interactive IR has focused on the conscious, as was mentioned in section 3.2 above. However, Harmon and Ballestaros (1997:422) contend that "to direct inquiry successfully, it is critical to generate and formulate initial questions that truly represent deeper cognitive needs residing in the unconscious." In their study, they found that greater elicitation of unconscious cognition through "information counselling" and a "programmed relaxation device" made for a "positive impact on the formulations and representation of research problems for the inquirer and the intermediary." Harmon and Ballestaros (1997:425) conclude that "unconscious cognitive processes form an integral part of any explanation of human information processing, and should be taken into account in the design of information retrieval systems."

Harmon and Ballestaros were inspired by Taylor's classic theory (1968) of information need development, which pointed to the importance of emerging, but still unexpressed information needs in the search process. The problem of eliciting from the user what it is they actually want to know has long been central to the reference interview. In Taylor's theory, the information seeker transits through four stages of information need, from a "visceral information need"

which may not be capable of being articulated, up to a fully developed need expressed as a statement suitable for the context of the information resources available. Hjørland (1997:163) has commented that Taylor is inaccurate: it is not actually the need that is changing but the awareness of the need, and so suggests an alternative term, "cognitive development," instead of "information need development." Hjørland (1997:165) goes on to critique, "Taylor's four phases thus represent a confusion of two different things: the development of the knowledge of the primary problem and the change in the information need as a consequence of this primary development." We have, in fact, already taken care to make the same distinction in our previous discussion of search goal redefinition, when developing a new model which takes account of both new knowledge and old knowledge brought to consciousness (see section 8.3). However, Hjørland's approach is collectivist (or social constructivist) and this leaves little room for an exploration of where the "development of the knowledge of the primary problem" might come from. As Harmon and Ballestaros propose, it would appear that a good deal of this development is due to contributions from the information seeker's individual unconscious, or as we will later label it, their *nonconscious.*

3.9 Use Studies of Real-Life IR Systems

The ways in which different types of user interact with IR systems has been the subject of several studies. However, a clear picture has not emerged, probably because user behaviour differs from system to system, and because the behaviour of user groups evolves as they become more system-savvy. Ballard (1994) revealed that the differences between patrons and staff were not always so great, although the latter were considerably more specific with their search terms, and employed some advanced searching techniques whereas the patrons did so rarely. Sloan (1991) found remote users less likely to perform subject searches; on the other hand, Snelson (1994) found search behaviours similar for remote and terminal users. In his study with a specialist system, Nicholas (1996) detects different kinds of searching carried out by end-user and intermediaries. End-users tended to be more selective at the citation list, being more aware of the "slanted" information they needed; end-users were also found to be more easily

satisfied, conducting shorter search sessions, and placing more emphasis on precision than on recall. In any case, the effect that search intermediaries have on end-users' information seeking needs more investigation.

With developments such as openURL, and the drive towards federated search functionality, we have now reached a stage where research into OPAC use and research into the use of other IR systems, may be more readily integrated. However, in the past, use studies were most commonly based on library catalogues, and we shall devote a section specifically to them.

3.10 Catalogue Use Studies

Research into catalogue use predates the OPAC, but most of the pre-OPAC studies reviewed by Hafter (1979) indicate relatively simplistic research and a focus on basic questions. However, two studies are worth noting here. Lipetz (1970) found that 57% of patrons would look up a book in order to identify appropriate headings for a further search – in other words, they performed "pearl" searches; and Frarey (1953) found that 50% of patrons used the catalogue to identify a shelf location for browsing, another kind of "pearl" search, and a result echoed by Hancock-Beaulieu (1990) with an OPAC, and a strategy the author also observed frequently during his own pilot study of a public library OPAC (chapter 4).

The research carried out on the use of OPACs is significant in quantity, although many studies do not "evaluate anything deeper than initial reaction and satisfaction" (O'Brien 1994:232). Many of the studies use transaction log analysis; some combine this approach with interviews or online surveys, while others use protocol analysis or mediation analysis. Several "before-and-after" studies have been carried out on OPACs and search success, although the latter is usually measured in the transaction log analyses in terms of non-zero hits, not-too-many hits, and/or valid search syntax and commands (Blecic et al. 1999). Thus "success" is based on strategy and ease-of-use more than on optimal content (Large & Beheshti 1997).

Beheshti (1997) has outlined the advances made with library OPACs over the past two decades. Three generations of OPACs can be discerned: a first generation based on limited computer power, with few access points; a second generation which was menu-driven, and allowed for Boolean searching; and a third generation with GUIs, probabilistic ranking, links to other OPACs (via Z39.50 etc.), and to and from the Web (often termed WebPACs). Most commercial OPACs, however, are still largely text-based, although some experimental OPACs feature graphical displays of collections and search results.

Commercial WebPACs have so far focused as much on cosmetic improvements, as on new functionality. Indeed, Beheshti (1997) considers that many Web-based catalogues fail to provide the richness of data that the traditional OPACs did. While they may display more presentable interfaces to users familiar with the search engines of the World Wide Web, they have been slow to develop aids for users in the formulation and reformulation of search strategy – few possess a full range of such features as automatic conflation/truncation, synonym searching, spell checking, best match techniques, relevance ranked output, graphical search trees displaying bibliographic relationships. Beheshti's observation (1996:181) remains valid: "transforming information needs into searchable queries for most users remains a difficult task."

Beheshti (1996:181) calls for further research in order to turn the experimental OPACs into a new generation of successful commercial ones:

> A better understanding of a user's 'inside' world of information and how it is organised is a prerequisite for the development of new systems that stimulate information worlds. The closer these worlds correspond to the 'inside' concepts of information and their internal relationships the more easily users will be able to navigate through them.

Such research has been slow in coming. Connaway, Budd and Kochtanek (1995), for instance, examine the kinds of search performed on an online catalogue, and call for OPACs to cater for "any natural language formulation of the user's needs," but they do not consider that the user's needs may not match his own formulation, that is, a user's intent may not be accurately reflected in the expression of his intent.

Much of the research into search behaviour on OPACs has been carried out on "second generation" catalogues. In recent years, researchers have turned to other types of IR system, in particular, Web search engines and other full-text systems, a reasonable development given the reality of usage. Although OPACs are still frequently used by library patrons, many other IR systems are now widely available to a networked public. Ray and Lang (1997) show how online access to indexing services has impacted on the use of library catalogues. Wallace (1997) likewise shows how periodical title searching on OPACs has increased as users toggle between online indexing/abstracting services and OPACs in pursuit of full-text articles. Web search engines are also squeezing out OPACs, a phenomenon highlighted at many library conferences.

As Beheshti (1997) argues, researchers must give more attention to the context of OPAC use, with so many other information outlets now available: further study is required to determine how users employ OPACs in their overall information seeking behaviour. As Spink and Beatty (1995) point out, the information seeking process often includes multiple search sessions on online catalogues and other databases as part of a task-oriented information seeking process (hence the importance of "search history" functionality).

3.11 Summary

User-oriented approaches are now quite well-established in information retrieval, influenced by the work produced by researchers of information seeking in general. User-system interaction is seen as a central, if not the central, subject of study, and this has led to more dynamic models of system use. In these models, search behaviour evolves at different levels, although the level of strategy is still often emphasized, at the expense of the higher levels of information goals and information wants.

The dynamic nature of information goals and wants, even in the context of a single search session, has nevertheless been indicated by researchers who have emphasized the *situational* nature of relevance judgements, which are, in effect, reflections of information wants, usually conceived in terms of search goals. The

importance of "situation" means that "not only planning but also feedback from the interaction influences the information seeking process" (Xie 1997:103).

In order for the model of user-system interaction to sit more comfortably in the wider context of information acquisition, the author is of the view that more attention needs to be given to information encountering, as described by information seeking research, during user-system interaction. Incidental information acquisition can play an important role in search behaviour if information goals are treated as revisable, and if such revision is based on dynamics of conscious and unconscious thought. Harter (1992), as well as Harmon and Ballestaros (1997), emphasizes how the cognitive state of the user may change during the course of a search session.

Robins (1997, 2000) and Dalrymple (1990) both show that query reformulations frequently occur during search sessions, and while a considerable amount of such reformulation may be strategic, the evidence suggests that some relates to goal redefinition. The traditional model of matching information needs with information available in the system may not apply if the user has an incomplete notion of what their need actually is. Although commercially available IR systems, even OPACs, are now more user-sensitive, Borgman's vision (1996:500) remains pertinent:

> We need to design systems that are based on behavioral models of how people ask questions. Such a design model could assist in the question-negotiating process, allowing the searcher to pursue multiple avenues of inquiry by entering fragments of the question, exploring vocabulary structures, capturing partial results, reformulating the search with the assistance of various specialized intelligent agents, retaining elements of a search for future sessions, and even transferring elements to other systems.

Chapter 4 PILOT STUDY

4.1 Introduction

The first research question to be addressed in this study, expressed as hypothesis 1 in chapter 2, is that users redefine their search goals during the process of searching bibliographic information systems on a significant number of occasions. Should there be evidence to the contrary, then answers to the other research questions discussed in chapter 2, would become less relevant. Thus it was important to test this primary hypothesis in a pilot study. It would also be useful to obtain a preliminary picture of the nature of search redefinition in order to assist in formulating specific hypotheses, and necessary to explore methodologies which might be used to investigate subordinate research questions. This pilot study is thus reported separately below, as it was essentially of an exploratory nature, necessary in order to identify a suitable methodology for the main analysis, which itself is detailed in chapter 5 (as is the piloting of that methodology).

The initial plan was to investigate search redefinition on standard library OPACs, and through this pilot study the author sought to find out whether this was practicable. It was hoped that certain OPACs with different types of record display could be compared in terms of the amounts of redefinition that they caused.

4.2 Methodology

4.2.1 Dedicated OPAC Set-Up

The OPAC dedicated for the pilot study needed to be a real-life OPAC, since the research aimed to investigate real-life search behaviour. The author received permission to use the OPACs of four libraries in Singapore, for research purposes. A set of OPACs in a large branch of the public library system was chosen, as they were used by a broad cross-section of society and provided

access to a bibliographic database representing a very large and general collection.

The public library system in Singapore had recently switched to a reasonably user-friendly, Web-based OPAC developed by the US library automation company, CARL. It is part of a suite of software (CARL Solution) and is called CARLweb. The public library system is run by the National Library Board of Singapore (NLB). The NLB catalogue also includes datasets representing the collections of the National Library, and the Singapore National Bibliography. Users of the OPAC at any of the branch libraries are able to search for materials irrespective of which branch libraries may hold them. The OPAC can be accessed remotely through the Internet, as well as at the libraries.

Although there is a transaction logging function available on this OPAC, it had not been switched on by the NLB technicians. The logger might have been switched on for the purposes of the author's research, but the function was found to be of limited value anyway. It logged the terms and type of query submitted by users, but could give no indication as to which terminal in which library (if any) had been used to enter a query, nor any output on the part of the system. It did not log any other actions taken by the user, such as record selection or the "next page" command.

A methodology that allowed for a more detailed analysis of search redefinition at the OPAC was therefore sought. It was decided to conduct pre- and post-search session interviews of users at the terminal, and for this data to be supplemented by capturing the users' actual interaction with the OPAC by means of ScreenCam software. ScreenCam is part of Lotus SmartSuite. It records the screen in real time, regardless of application being run, acting as an internal video camera.

Permission was granted for one PC, of those dedicated to OPAC use, to be installed with the ScreenCam software. There was in fact a choice of position for the ScreenCam OPAC: either as one of two OPACs with seats, or as one of three OPACs without seats. A "sitting OPAC" was first selected since it was thought

that generally more "serious" searching would be carried out there. However, this OPAC proved to be used much less than the opposite bank of "stand-up OPACs," and so the ScreenCam software was switched to one of the latter.

The PCs were networked in order to link to the OPAC server; for security reasons, the author had to rely on the technicians to access the PC's hard disk, onto which the software was installed (it had to be run from the hard disk). It was soon established that the files of recordings were going to be very large – even one minute's worth could easily produce a file 1.5 MB in size. The library's PCs were not particularly new and did not have enough space on their hard disks for such files. The problem was solved by setting up so that the files would be created on a Zip drive. A Jaz drive (from Iomega) with a 2 GB capacity was duly connected to the PC. Each recording session would produce a separate video file – those containing no applicable searches were later deleted.

Although the PC was supposed to have enough RAM to run ScreenCam, over extended periods of time it was found not cope very well. In other words, it had a tendency to hang, which resulted in the loss of all data for that recording. It was thus imprudent to run ScreenCam continuously for more than ten minutes at a time. Since re-starting the recording took approximately one minute, and had to be done at the OPAC, in public view, this was a significant drawback. Often this would be necessary even though no applicable search session had been recorded during the ten-minute period; and often the recording had to be re-started when there were several potential users nearby – they could be queuing to use one of the OPACs, including the one set-up with ScreenCam.

Unfortunately, not many search sessions performed at the terminal could be used during the course of a typical day at the OPAC, for users were often unwilling to participate in the interviews, either because they did not wish to waste time, or because they were shy of relating their OPAC use. At certain times of day, the OPACs were, in any case, in little demand. During such "quiet" periods, the number of interviewed sessions was about one per hour.

For a more efficient use of research time at the OPAC, interviewing was limited to the busier searching periods: the latter part of the afternoons and evenings, Tuesdays to Fridays, and Saturday and Sunday afternoons. Since the author (who worked full-time and some evenings) was unable to attend during many of these hours, he employed two research assistants for the purposes of this pilot study. The two assistants were students of a local information studies pre-university diploma course and had worked with the author on previous projects.

The author worked with the assistants at the library during two preliminary sessions, establishing consistency of interviewing technique. During the following two weeks, the assistants spent a total of 30 hours at the OPAC, collecting 106 search session interviews. A search session was defined in this pilot study (but not in the main study) as one based on a particular search goal – thus the same user might perform more than one "search session" during one "OPAC session."

The interview responses were noted in writing by the research assistants on the interview schedule; each set of responses was linked to the serial number of the relevant ScreenCam recording.

The users were not informed of the recording, because in the author's judgement, this might otherwise prejudice the findings of the study – users might search differently, or respond to the interview questions differently. This raised an ethical question which was discussed with the University, as well as with the host library. It was considered a reasonable approach given that the interviews linked to the recordings were conducted anonymously, with persons "off the street," who were not asked for any personal details, and that there was no way in which the author, not present at the time of the interviews, could identify the persons behind recordings. ScreenCam recorded only the OPAC screen, with the OPAC being locked down on the interface, and no other images, such as of the user, nor any sound. The OPAC sessions revealed no personal details of the users, with no logins being needed. Although the research assistants would re-set the ScreenCam recorder, only the author was able to play back the relevant search sessions (only the recorder was on the library PC, not the player); the ScreenCam

data was stored securely on the Jaz disks, to which only the author had access. Thus while the students had contact with the users, they did not view the recordings; and while the author viewed the recordings, he had no contact with the users.

Express permission for both the interviews and the screen recording was given by the manager of Tampines Regional Library, Singapore, the public library in which the study took place. It was recognized that transaction logging routinely occurs without users' knowledge -- and certainly without their express permission -- on library management systems, as would be the case on the OCLC FirstSearch service, the logs of which were used in the main analysis (see chapter 5). The critical provision is for anonymity to be guaranteed – the FirstSearch logs were de-identified by OCLC, before they reached the author.

Unfortunately, upon detailed inspection of the ScreenCam data, it was discovered that many of the search sessions had not been recorded clearly – the movement of the cursor would progressively blur the screen. Thus only eleven of the 106 search sessions were analysed using the ScreenCam data as well as that of the interviews.

4.2.2 Interviews

4.2.2.1 Interview Questions

The primary purpose of the pre-search session interview was to establish the intended search goal or goals of the user. Firstly, the potential interviewee was identified as someone queuing for or walking up to the bank of OPACs (and who had not witnessed the re-starting of the ScreenCam software). He or she was asked if they minded answering a few questions before and after their OPAC session. They were informed that interviews were anonymous, and that information was to be treated with appropriate confidentiality. If the user agreed to participate, then the following two pre-search session questions were asked. (If they did not agree, then no interview was conducted, and ScreenCam was not used.)

1. How often do you use the NLB OPAC?

2. What will you be looking for in this search session?

The user was then invited to conduct their search session(s) on the ScreenCam OPAC. An assumption made here was that the user was unaware that their searching was being recorded by ScreenCam – the researchers only selected potential interviewees who they did not think had seen the re-starting of the recording. The aim was to avoid the possibility of search behaviour bias due to the self-consciousness that awareness of ScreenCam might incur.

During the search session, the research assistants would stand aside, well away from the OPAC, so that they could not see, and could be seen not to be able to see, the details of the user's searching. At times users actually invited the researchers back to assist in their searching; naturally the researchers refrained from doing so. However, this in itself might have caused a bias in some cases, since the interviewees who had been denied assistance might have developed a different attitude compared with that of the interviewees who had not sought it, which in turn might have affected the post-search interview.

The OPAC interface offered only a "basic" search interface, such that search formulations were automatically treated as keywords (unless the user had prior knowledge of the system's commands). Post-results, users were however invited to specify particular fields to refine the search.

Immediately after the user had completed their OPAC session (which might represent more than one search goal as well as any redefined ones), the post-search interview was conducted, with the following questions being asked.

3. Were you satisfied with the results you got from the OPAC?

4. Did you find what you were looking for? What did you actually end up finding?

5. Are you now going to go and look for any of the items you found on the OPAC? Which ones? If none, is it because they are not available at this library?

6. During your searching, did you think of things you wanted, that you hadn't thought of in the first place? If so, what?

7. Did the results you got give you more ideas as to what exactly it was you wanted? If so, please explain.

The user was then thanked for their time.

4.2.2.2 Interpretation of Interview Responses

All interview questions were open ones, since the researcher wished to minimise his influence on the responses and, in the case of the pre-session questions, on the searching.

The first three post-session questions cover user *satisfaction* and goal *realisation.* Whilst the three questions may be linked, they are not necessarily so. Users may have been "satisfied" with the results the OPAC produced, without these results

being positive ones – they may be satisfied in the sense that they are satisfied the OPAC has accurately reflected a lack of items in the collection relevant to their search. However, it is likely that many users would not interpret the question in this way, and instead interpret it in terms of information goal realisation.

Nevertheless, the question is still different from that of the second post-session question, since a user may have realised a redefined goal rather than what he was originally looking for: because users had already described their initial search goal to the researcher, they generally answered the first part of Question 4 in terms of this initial search goal, so that they might not have found what they were looking for, but were satisfied with results, overall.

Furthermore, a user may not be satisfied with the results even if he found one or more relevant ones – they might be only partially relevant or the items might not be available at the library. The latter scenario demonstrates the important difference between *relevance* and *utility* (Cooper 1973). Questions 4 and 5 were also designed to draw out this distinction, and in some instances they did so. Question 5 may not be answered positively even though Question 4 was: the user may have found something at the OPAC of relevance, but could not follow up by retrieving the item because of its unavailability. However, where the response to Question 5 is positive, then whether the items the user was to seek represents a departure from the initial search goal had to be judged by the researcher according to the user's answer to the pre-session goal as stated in Question 2. Even if it does not represent a deviation, this does not necessarily mean that no redefinition has taken place – instead, it might mean that the more relevant items, according to a redefined search goal, were simply unavailable.

The last two post-session questions both ask about search redefinition directly, in different ways. A positive answer in both cases was taken as strong evidence that the user considered redefinition to have taken place; a positive answer to only one question was taken as mixed and insufficient evidence. Redefinition may have taken place at the OPAC without the redefined goal having been fulfilled, so these questions were again not quite the same as Questions 3-5. Questions 6

and 7 covered *any* redefinition, regardless of the extent to which the redefined search goals had been realised.

Six possible scenarios for a search session, defined by a single, specific search goal, were identified as follows:-

a) no results relevant to initial search goal found, user ends session

b) result(s) relevant to initial search goal found, user ends session

c) no results relevant to initial search goal found, user redefines goal and continues; eventually relevant results found, user ends session

d) no results relevant to initial search goal found, user redefines goal and continues; relevant results still not found, user ends session

e) result(s) relevant to initial search goal found, but user redefines relevance and continues searching; more relevant results found, user ends session

f) result(s) relevant to initial search goal found, but user redefines relevance and continues searching; no more relevant results found, user ends session.

The interview responses were analysed to determine which of the six above scenarios each search session represented, according to the following criteria.

If the response to Question 4 was that the user had not found what they were looking for, and this was not contradicted by the response to Question 5 in the light of the response to Question 2, *and* if Questions 6 and 7 were both answered in the negative, then the session was categorised as an instance of scenario (a).

If the response to Question 4 was that the user had found what they were looking for, or would be if available at the library, and this was not contradicted by the

response to Question 5, *and* if Questions 6 and 7 were both answered in the negative, then the session was categorised as an instance of scenario (b).

If the response to Question 4 was that the user had not found what they were looking for, and this was not contradicted by the response to Question 5 in the light of the response to Question 2, *but* if Questions 6 and 7 were both answered in the affirmative, then the session was categorised as an instance of scenario (d).

If the response to Question 4 was that the user had found items of *relevance,* or would be if available at the library, but that these items do not follow directly from the response to Question 2, and this was not contradicted by the response to Question 5, *and* if Questions 6 and 7 were both answered in the affirmative, then the session was categorised as an instance of scenario (c) or (e).

If the response to Question 4 was that the user had found items of *relevance,* or would be if available at the library, but that these items do follow directly from the response to Question 2, and this was not contradicted by the response to Question 5, *and* if Questions 6 and 7 were both answered in the affirmative, then the session was categorised as an instance of scenario (f).

Those sets of responses that did not meet the criteria for any of the six were deemed to exhibit mixed evidence, and as such were set aside.

An additional, more detailed analysis was carried out with eleven sessions recorded by ScreenCam. The analysis of the interview responses was followed up by analysis of the search expressions entered into the OPAC and the selections from the citation lists.

4.3 Findings

4.3.1 Findings from Interview Data

The pre-search interview revealed different levels of OPAC use on the part of the public library's patrons, as expected, from first time, to almost every day. The

initial search goal expressed by the user before interaction with the OPAC also varied from broad subjects to specific items. Out of the 106 sessions, answers to Question 2 indicated that 81 (76%) had subject goals, 15 (14%) had specific item, work or series goals, and 10 (9%) were for specific authors.

Out of 106 sessions, 4 sets of responses were set aside due to their self-contradictory or unclear content. Of the remaining 102 sessions, in 64 cases (63%), no search goal redefinition was identified. Of these, 38 sessions (59%) were at least partially successful – session category (b) above. (The definition of successful here is that an item is found that might fulfill, at least to some extent, the search goal.)

In 38 sessions (37%), some search goal redefinition was reported. Of these, in 28 cases (74%) this redefinition led to at least partial success – i.e. session category (c) or (e) above.

The proportion of subject search sessions where redefinition was considered by the user to have occurred was 42%. Given an overall proportion of 37%, this suggests that subject searches might be particularly liable to redefinition, as one would perhaps expect.

Users in 40 (82%) of 49 sessions involving articulated redefinition expressed satisfaction with their results, contrasting with only 34 (52%) of 66 sessions not involving articulated redefinition (one interview being discarded due to an unclear response).

The primary *reasons* for search redefinition articulated by the user, for each session involving a subject goal, are shown in table 4.1 below.

Reason	*f*
Ideas from titles	17
"Self-inspiration" / spurred by initial failure	13
Ideas from subject headings	2
No reason given	2
	34

TABLE 4.1 PRIMARY CAUSES OF SUBJECT SEARCH REDEFINITION AS ARTICULATED IN POST-SESSION INTERVIEWS

The explanation type, "It just came to me," termed above as "self-inspiration," was hard to distinguish from some statements indicating "Hobson's choice," where given "poor" results, user felt themselves to have little choice but to think of a new expression. The articulations were sometimes quite vague, and the author suspected that the "reason" might in some cases represent a lack of articulation at the pre-session interview more than actual redefinition, and an inaccurate response to Questions 6 and 7.

Nevertheless, in another 19 of the interviews specific reasons relating to the displayed OPAC content were given, and at least in these cases, the author considered the articulation of redefinition highly likely to reflect actual redefinition.

4.3.2 Additional Findings via ScreenCam

Several additional types of data were derived from the ScreenCam recordings: the queries entered into the OPAC; the citations returned and full records displayed by the user; whether the full records selected were paged down in order to display their items' availability; and, if so, whether they were indeed available.

The data derived from the eleven ScreenCam recordings is given in table 4.2 below. Although the data represents only a small sample of OPAC sessions, it was used to shed further light on the feasibility or otherwise of a full-scale analysis of search goal redefinition.

TABLE 4.2 SCREENCAM-RECORDED SEARCH SESSIONS

Search session	Search concept(s)	Search input	No. of citations displayed*	No. of records displayed	Citation no. of record	Selected/ Available
1	Web page design	Web page design	14/32	4		
					8	S/NA
					13	S/NA
					12	S/NA
					4	S/NA
		Microsoft FrontPage *(from last record's other title)*	14/126	4		
					2	S/NA
					1	S/NA
					5	S/AE
					11	S/NA
2	Indonesian food/cookery	Cookery	14/6917	1		
					13	NS
		Indonesian food	28/42	2		
					2	S/NA
					10	S/AE
3	Fictions books called Sleepers	Sleepers	14/26	1		
					8	S/A
		Replica	7/34	1		
					1	S/NA

4	Foreign talent / national education	Foreign talent	49 *(browsing adjacent titles)*	1	n/a	NS
		National education	7/1016	0		
5	Books on Visual Basic C++	Visual Basic C++	15/15	1		
					3	S/NA
		Visual C++ *(from record title)*	5/5	1		
					4	S/AE
		Microsoft Visual	28/688	1		
					28	S/AE
6	Microcontroller technology	Microcontroller technology	3/3	1		
					1	NS
7	Witchcraft	Witchcraft	7/179	1		
					2	S/NA
		Witchcraft History Sources *(from record's subject headings)*	3/3	2		
					3	S/NA
					1	S/NA

8	Astronomy	Astronomy	14/1384	2		
					2	S/AE
					3	NS
		Planets	21/866	3		
					12	S/AE
					14	S/AE
					20	S/AE
9	First World War	First World War	21/266	1		
					19	S/AE
10	Fiction books called Chinese Cinderella	Chinese cinderella	0			
		Lu xiao feng	7/58	1		
					3	S/NA
		Chinese cinderella	7/10	2		
					1	S/NA
					2	S/AE
11	Filippo Brunelleschi	Filippo Brunelleschi	2/2	1		
					1	S/NA
		Donatello	7/7	1		
					1	S/AE

S=Selected as indicated by paging down for call number & availability
NS=Not selected as indicated by paging down for call number & availability
A=Available in the branch library
AE=Not available at the branch library, but available at one or more other branches
NA=Not available at any branches (e.g. borrowed or on order)

*This ratio (number of postings displayed / number of hits) was used via a different system for the study of persistence described by Wiberley, Daugherty & Danowski (1995). A ratio which shows a bolder investigation, i.e. the number of *fuller* records displayed / number of postings displayed, can be derived from the numbers in the fifth column and the numerators in the fourth.

Note: none of the sessions above included any reservations for unavailable items

First, comparison was made between the search goal expressed in the pre-session interview and that expressed at the OPAC. In six out of eleven cases, the expressions were identical; in three further cases, only the format (books) was omitted from the OPAC formulation, which is likely to be either for strategic considerations (it was considered unnecessary to include this aspect in order to obtain the desired results) or because there was no obvious limiting function offered to the user. In the final two cases, two related concepts were expressed during the pre-search interview, and were expressed in two separate queries at the OPAC. Thus it was found that the users' articulation of their search goals mirrored very closely that expressed in their queries, which bodes well for an analysis reliant on either type of data.

Second, association between post-session articulation of search redefinition and query reformulation indicating possible search redefinition, was likewise examined. Of the eleven ScreenCam sessions, in two cases, only one query was entered; in two other cases, the two queries entered aggregated exactly to the search goal articulated during the pre-search interview. In the seven remaining sessions, additional concepts were introduced in subsequent queries or existing concepts were omitted – in the view of the author, at least. These observations are compared with the responses to the post-session questions concerning redefinition in table 4.3 below (Y=Yes, N=No).

Search session	OPAC expressions diverged from stated goal?	User considers search goal redefined in session?
1	Y	Y
2	N	N
3	Y	N
4	N	Y
5	Y	Y
6	N	Y
7	Y	Y
8	Y	N
9	N	N
10	Y	N
11	Y	N

TABLE 4.3 EXPRESSIONS OF REDEFINITION

From the above table, we see that only five out of the eleven sessions produced matching results. This result is discussed further in section 4.4.2 below.

Some of the users' statements that record displays contributed to search redefinition were supported by the ScreenCam evidence, the video showing the relevant displayed content, in some cases with the cursor moving over the words or the words being copied and pasted for the new query.

Finally, the ScreenCam recordings indicated, as did the interview responses, that in some cases of redefinition, the "inspiring" content was at the citation-list level rather than at the full-record level. However in other cases, the ScreenCam recordings showed that redefinition was very probably induced by an element of the full record, such as a subject heading, which was not included in the brief citation.

4.4 Conclusions

4.4.1 Conclusions concerning Search Redefinition

The extent of redefinition indicated in the pilot study was considered significant such that its nature and cause was a worthwhile topic of investigation. Of the 81 subject search sessions, 19 (23%) were very likely to have included some goal redefinition, given that in these cases a specific cause of redefinition, relating to user-system interaction, was articulated by the user. In a further 15 cases, users said that redefinition occurred, although a specific cause was not stated. If system-induced redefinition took place in all of these search sessions, the proportion rises to 42%. Perhaps the *actual* proportion lies somewhere in between 23% and 42%.

From the pilot study and its findings, several points about the possible nature of search redefinition were noted. First, it is considered likely that some redefinition does *not* result in additional queries being entered into the system, at least not during the same session. This kind of redefinition is termed here *weak redefinition.* Redefinition that does result in one or more subsequent queries is

termed here *strong redefinition.* It appeared from the pilot study that weak redefinition would be hard to detect and thus hard to measure. Thus it was hoped that enough redefinition occurring in real-life IR situations was of the *strong* kind, in order for this to be readily detectable and measurable.

The author also realised that search redefinition can be negative, in the sense of limiting, as well as positive, in the sense of transformational or expanding. For example, one of the users made it clear in his post-session interview (ScreenCam session 6) that they "realised" the books in the results set were too old. Users may not think of particular parameters at the outset, but identify them later when details in citations breach these parameters.

It was also recognised that during many of the OPAC sessions users examined (displayed) only a small number of full records, and that a large amount of redefinition may be based on the brief citation lists. Nevertheless at least *some* redefinition was derived from full-record displays.

Queries may be reformulated due to redefinition, but also due to search failure. Several of the responses to Questions 6 and 7 indicated this in fact, as did a couple of the ScreenCam recordings the author observed. For example, the user would find two or three apparently relevant records only to find that they were unavailable or not in that branch library; the user would examine these records more than once, scrolling down seemingly to confirm their unavailability at the branch; finally, they would copy a subject heading or term from a title, and enter a new search; if they found the first one or two titles in the new results set available, they would then halt their session and presumably follow up on those call numbers. The author considered a likely interpretation of this behaviour to be that the user had been frustrated by the initial items' unavailability and tried a second query not necessarily because they had determined a more accurate expression of their search goal, but because they simply wanted some material that would at least partially satisfy their information need.

Query reformulation due to search failure may be purely strategic in nature, or it may nevertheless incorporate an element of goal redefinition: users may be more

alert to new angles or aspects because they consider the items they have retrieved up to that point to be insufficient (as opposed to unavailable); if users had been "happier" with their search results, they might not have been so alert.

A fundamental premise in this research is that search redefinition generally increases the satisfaction, or at least potential satisfaction, of the user's information need/want. In this way, redefinition is to be welcomed. While this premise is very difficult to test, an indication that it might be true, at least in some IR situations, was provided by the pilot study: a much higher percentage of users in sessions involving articulated redefinition expressed satisfaction with their results compared with those in sessions not involving articulated redefinition.

4.4.2 Conclusions concerning Methodology

Although in theory the ScreenCam software could be used as an alternative to an orthodox transaction logger, it had not proven a very practical tool for the study of real-life OPAC use. The truth is that it was never intended for such a purpose; rather, it was meant to be used to record brief clips for inclusion in software tutorials and the like.

In the pilot study, the user's articulation was the primary indicator of their search goal. In the main study, the query entered into the system was to be the indicator, with an assumption made that this representation of the search goal approximated the actual search goal such that any discrepancies did not have a significant effect on the overall results. The similarity of articulated and entered expressions in the pilot study provides evidence to support this assumption.

A much less convincing degree of association was found between articulated and entered redefinition. However, on reflection, one would expect this to be the case, even though the pre-interview expressions were all represented at the OPAC. First, there may well be times when the whole search goal fails to be *articulated* in the interview, but is nevertheless in the mind of the user at the start

of the session. Second, some users might not be prepared to admit their change of mind during the post-session interview.

There are, in fact, at least two other important reasons which may contribute to a lack of association, unrelated to discrepancies between user articulation and system use.

First, as we have already noted, some of the query revisions during a session might occur out of *strategic* considerations and not due to redefinition. One or more of the sessions 3, 8, 10 and 11 in table 4.3 might have been instances of this. Certainly the revised query in session 10 could well have been intended as simply another way of obtaining a work-specific search. This has important methodological implications: such cases need to be distinguished from those where redefinition has occurred, *or* an assumption needs to be made about search strategy not affecting the overall findings of a study on redefinition.

Second, a user may redefine their search goal, but this might not necessarily result in another search – what was termed in the previous section as *weak redefinition.* For instance, upon examining a results set, a user may realise that some citations may be particularly relevant because of certain features they had not thought of at the time of inputting the search, but does not consider it necessary to perform another search, because the citations they have come across suffice. Sessions 4 and 6 in the table 4.3 may have been instances of this scenario – probably a more likely explanation for the mismatch of data than that the user falsely claimed goal redefinition.

It appeared much more difficult to investigate weak redefinition than strong redefinition. The former may not be demonstrated in *any* of a user's actions on the system. It is quite possible that they will not display a full record which indicates the revised goal, or the revised aspect of it, subsequent to the conceptualisation of the revised goal. Even if a full record display, or a cursor movement, does provide a clue, it is highly questionable whether the researcher would be able to consistently identify it. Thus weak redefinition could really only be investigated through its articulation by the user.

Strong redefinition, on the other hand, was identified through users' responses to the post-session questions, but also through observation of screen recordings and confirmed through reference to the pre-session interview. While weak definition might or might not be recognised by the user in a post-session interview, in most cases where a new query representing strong redefinition was performed, users articulated redefinition more clearly. Thus the pilot study had shown that it might be more reliable to focus on strong redefinition.

Chapter 5 METHODOLOGY

5.1 Choice of Methodology

Analysis of search behaviour occurring during interaction with an IR system can be (and has been) based on a range of methodologies. The pilot study described in chapter 4, was based on a combination of two methods – a form of transaction log analysis and the interview survey – and there are others which might be usefully employed to investigate the research question detailed in chapter 3. Moreover, there are several ways in which each of these methodologies might be executed.

The "behaviour" we are investigating actually takes place in the searcher's head – search goal redefinition is a construct. The author takes the view that this construct may be identified through observation of certain external phenomena caused by the searcher's actions and/or expressions. Studies of the "cognitive" aspect of information retrieval rely on this view, of course. The question is which external phenomena to observe: protocol analyses and interviews are based on direct, aural communications; questionnaires (including online ones) may be based on direct, written communications; mediated search studies on indirect, aural communications; and transaction log analyses on indirect, system-oriented actions.

In the first part of this chapter the possible methodologies – protocol analyses, interviews and questionnaire surveys, mediated search studies, and transaction log analyses – will be reviewed in light of the research question.

5.1.1 Protocol Analysis

In protocol analysis, searchers are asked to relate the thoughts they had during a search session by "thinking aloud." They may do this either at the time of the actual search session, or afterwards, through a replay. The methodology has been used in several IR experiments (Powell 1999:111). Often the searcher's verbal output is recorded on audiotape, and then transcribed and analysed in relation to

aspects of the searcher-system interaction. In some cases, the searcher is videotaped; however, non-verbal cues may be deceptive and have rarely been noted in IR studies.

"Thinking aloud" is arguably the most direct representation of a searcher's thoughts. Unlike most interview and questionnaire surveys, it takes place during the search session and its accuracy is thus not compromised by the fallacy of human memory. Protocol analysis can be used in combination with other methods, such as the post-session interview.

The directness of the methodology is also a drawback: it is the most intrusive of techniques, with the possible exception of an interview conducted during the search session. Protocol analysis necessitates a self-conscious searcher. More details of a searchers' thoughts may be recorded, but they might not always be those that the searchers would have had in a regular searching session; nor would the protocol necessarily represent *all* their thoughts – some may be glossed over (consciously or otherwise), while the searchers may be unable to verbalise others (perhaps in some cases due to competing, simultaneous thoughts). Researchers should not assume that the transcription tells the whole story.

5.1.2 Interviews and Questionnaires

Another approach is to elicit the pertinent thoughts of searchers by asking them questions before, during, or after the search session. In order to encourage participation, structured interviews may be favoured, instead of requesting the participants to complete questionnaires themselves, although online questions may be produced that are reasonably easy and not too time-consuming for participants to answer.

An advantage of direct questioning is that feedback is focused (or should be) on the issues under investigation. For some types of investigation, post-session reflection on the part of the searcher may be helpful, although it also represents the danger of "elaboration." A pre-session interview or questionnaire may be an effective way of eliciting information critical in an analysis of subsequent search

behaviour; at the same time, pre-session questions and answers may influence the subsequent search behaviour.

An interviewer's presence during a search session may also influence behaviour, while even more obtrusivity results when the interviewer interrupts searching with questions. On the other hand, an online interface that skillfully weaves in questions as a search session progresses may represent a much less obtrusive technique and garner immediate feedback, which may be more reliable than post-session feedback. For instance, questions relating to whether a new search is the product of a reassessment of information goals, could be programmed to appear on the screen immediately after the searcher has clicked the "Search" button. The researcher might be required to introduce the programme to the potential participant, but he could then leave searchers to conduct their sessions in peace.

Nevertheless, there is a serious danger of pre-session or intra-session questions *leading* the searcher. A question can be formed in a relatively neutral way – for example, "Why did you perform another search?" instead of, "Have you changed your mind on what you are searching for?" But the very fact that the searcher has been asked to reflect on this topic may itself influence the way the session continues (or does not continue).

5.1.3 Mediated Search Studies

A mix of protocol analysis and interview elements is exhibited in another methodology that some researchers favour. An extended interview is carried out during the search, but with a difference: the participant is not doing the searching; rather, they are supposed to be *directing* the searching, through a trained mediator. This arrangement is based on a mediated search service. It allows for the gathering of data on users' reactions to output as well as formulations of input, and can be preceded by a reference interview – a detailed identification of the search goal.

This methodology is sometimes referred to as dialogue analysis, and is still carried out by researchers (e.g. Robins 1997). The interaction between end-user

and system is related through the dialogue with the search intermediary, as opposed to the monologue on which protocol analysis is based. Dialogue analysis may be regarded as relatively unobstrusive. The searcher may even by unaware of the recording of the dialogue (for analytical purposes) until after the event (when permission for its use is sought).

While mediation reduces the variable of searching expertise – all participants interact with the system through the same trained mediator – there is again the issue of leading questions to consider. The mediator has to be careful with their contribution to the mediator-participant dialogue: the participant should be the one who takes the initiative when discussing search goals; yet sometimes strategic suggestions also touch on the information problem/task. The very fact that the mediator is present is likely to make a difference: unmediated searches are likely to produce different search behaviour to mediated ones.

5.1.4 Transaction Log Analysis

Transaction logging is perhaps the least obtrusive technique for collecting data on search behaviour. It involves the automatic recording of searchers' input to a system. While transaction logs have often been analysed to identify the extent and nature of *misuse* of information systems, particularly OPACs, they may, if sufficiently detailed, provide the basis for more sophisticated analyses, for they show the expressions of a searchers' thoughts, albeit expressions constrained by the parameters of a system's inputting language. All expressions of thoughts are constrained by language.

Transaction log analysis (TLA) may be based on data collected totally unobtrusively, if users are not aware of the recording. Some ethical considerations need to be resolved here, but when users search on a system anonymously, or the logs are not identifiable, then disclosure of recording may not be necessary. Many OPAC systems routinely log searching, without users being conscious of this practice.

This research examines users' reaction to system output. Thus a transaction logger which records both input and output would be ideal; unfortunately, not many loggers do. Most loggers on OPACs and other information systems record only input. While protocols, discourses and question responses might give clues about output, they are unlikely to paint the full picture. Another technique to complement input logs might be for an external video camera to record the computer screen that the searcher sees, although this may be rather off-putting. Alternatively, researchers have sometimes replicated search sessions by re-inputting the users' queries shortly after their sessions, before the information system has changed in such a way that would affect output (Peters 1993). Such replication may even be more insightful than watching screen replays, since it induces participant observation: "visually analyzing a transaction log quickly enables researchers to put themselves in the users' place, and they begin thinking as the searcher must have thought at the time of the search" (Wallace 1993).

Finally, there is another means of recording both input and output that goes beyond *transaction,* but can be used as the basis of a TLA. This technique uses an internal *screen* recorder, such as Lotus ScreenCam (see chapter 4). Again, it may be executed totally unobtrusively. Screen loggers provide a fuller picture of user-system interaction and can be employed independently of the system. One problem is that they tend to consume a considerable amount of computer memory, and are normally not designed for extensive observation of user-system interaction. Before the advent of ScreenCam, just one screen recorder was mentioned by Flaherty (1993) in her report on the equipment available to TLA researchers, namely OLIVE, originally developed by City University for its experimental Okapi system. Although OLIVE was accommodated as a piece of front-end software by several commercial library systems, it fell away as these systems were upgraded.

TLA may be used in conjunction with other methodologies, or on its own. Some researchers have argued that ideally TLA findings should be verified through other methodologies, but this also applies to the other methodologies outlined above. For example, although interview or questionnaire data may more directly reflect thought processes, it may be less reliable as data about actual behaviour.

As Wallace (1993:240) argues,

> users' perceptions about what they do often do not match what they actually have done ... One of the problems with surveys is that they are really more tests of the users' memories, or the impressions that they wish to create, than accounts of actual behaviour.

5.1.5 Selection

The pilot study demonstrated practical problems surrounding the application of two of the methodologies outlined above, namely the interview survey and transaction log analysis. The author wished to observe real-life searching, yet this presented the challenge of recruitment. The active participation of "persons off the street" is not so easily secured, when a search session is only a means to an end and many users consider their interaction with an OPAC to be a fairly mundane affair. Often it is difficult to retain users for post-session interviews – they usually want to follow up their searching and not to spend time answering questions. An online survey may also be too tedious to complete. A brief interview might be palatable, but whether the researcher thus receives a detailed account of what actually went on in the searcher's mind, is questionable. Will the searcher be able to recall the reasons for a new search? They may have been only partially conscious of some of the reasons even at the time the decision was made. On the other hand, one does not want to prompt the participant with a leading question. The longer and more involved the search session, the less accurate may be the post-session answers. A participant's memory may well be improved if the session is played back to them, but fewer real-life searchers are willing to sit down and talk through a whole play-back.

Pre-session interviews or questionnaires (including online ones) may be more obtainable than post-session ones, since the interview or questionnaire may be positioned between the searcher and the system, such that there is a "captive audience," in the sense that access to the system is conditional on completion of the questions. However, the author remained unconvinced of the pilot study's pre-session interview. The question, "What are you going to be searching for?" is the obvious one to ask, but also the most dangerous – it encourages the searcher

to reflect on what the search goal is, reflection that might not otherwise have occurred, at least not until during the searching.

Whether or not pre-session interviews or questionnaires are used, the core data needs to be derived from the search session itself. Given the subtle nature of the subject of this research – search goal redefinition – and the impact of a large number of other variables involved in real-life search behaviour, the author considered it necessary to obtain a large sample of data. The amount and quality of the data obtainable through the post-session interviews as described in the pilot study, was not considered adequate.

The use of Lotus ScreenCam in the pilot study had also proven problematic, for practical reasons. When running for extended periods of time, or on complex operating system configurations, it was found to be unstable and unreliable. While ScreenCam can run on most versions of Windows, it is not designed for recording prolonged interface activity, but for the creation of two- or three-minute instructional or promotional clips. Very large amounts of memory were also required – most of the personal computers made available to the author in the library for the pilot study (see chapter 4) simply did not have the capacity.

In real-life situations, it would be preferable to keep the recording software running continuously, but this was not practicable with ScreenCam. The amount of time involved in re-starting the recording before a session was significant (usually taking over a minute) and this may well have deterred potential participants, or raised doubts in their minds about the integrity of the system. An alternative strategy would have been for the researcher to disclose the nature of the software to the participant, but this would raise the issue of obtrusivity, as well as that of recruitment. The library users in Singapore appeared to be uncomfortable with the notion of having their search sessions recorded, however innocent their searches might be.

Unfortunately, ScreenCam was the only commercial internal screen recorder available, to the author's knowledge, at the time of investigation. More recent entrants into the screen capture video market, such as "HyperCam" and

"Camtasia," are also not designed for extensive observation sessions, although they may prove to be more robust.

Given these problems with the methods used in the pilot study, other possible techniques were revisited and further explored. The external recording of user-system interaction by means of a video camera could be set up, but the presence of video cameras in public settings would deter volunteers, affect sampling, and probably influence search behaviour. Further, a video camera may not be permitted – the use of video cameras, even if they are intended only to film OPAC screens, would be something very hard for library managers in Singapore to accommodate.

Protocol analysis could be applied to ScreenCam recordings, since an audio option to capture external sounds (through a microphone) is available in the software. However, this would create even larger files and compound the software's practical limitations. A separate audio recording system could be set up at an OPAC terminal, but enlisting users "off the street" to perform recorded, talk-aloud sessions in a public space would be even more difficult than to obtain responses to interview questions. The talk-aloud sessions might be recorded in a private space (such as a closed room in the library), but in the author's library context, users – whether students, staff, or the general public – were generally reluctant to "think aloud" in any circumstances. Apart from their inhibition when faced with a microphone, there was normally a wariness of such intrusions into what is considered a private affair, especially when such intrusions are to be recorded. As a result, the number of willing participants at the academic and public libraries available to the author for such research was extremely low – a rate of fewer than ten per day, it was estimated. (The author draws the ethical line at using psychological tactics, such as playing upon a staff-student relationship, in order to persuade searchers to participate.)

If the research question could have been addressed experimentally, in the laboratory, using artificial information-seeking scenarios, then willing searchers might have been found more easily; but this research asks something about real

life, attempting to quantify and analyse *real-life* redefinition (which cannot be very easily simulated in any case).

One could study the searching of particular, willing individuals on a continuous basis, as and when they conducted searches (perhaps on a specific machine), but this would create sampling issues, and arranging for the *ad hoc* recording of their verbal output would also be difficult at a practical level. Perhaps an automatic recording device could be triggered upon the start of a search session on a particular machine (such as their office computer), but this requires quite sophisticated, customised technology, which was not available to the author.

The possibility of online questions, before, during or after the search session, was also considered, and it might have been that such questions would have obtained a more willing response than oral interviews. However, this approach also required customized technology, as well as permission from the library. It might also require access to the IR system's programming, certainly in the case of an integrated mode of presentation, and the system proprietor's permission. With respect to real-life, commercial IR systems, such access and permission is hard to come by, and was not available to the author; nor was the necessary programming.

Discourse analysis based on mediated searching was also considered. Indeed, if mediated searching was still the typical form of online information searching, then this might have been pursued. However, in the Singapore library context and outside in the wider information systems context, this was deemed no longer to be the case. Most modern systems possess end-user-friendly interfaces and the trend appears to be towards fewer and more limited mediated search services (Tenopir 2001).

It might be argued that mediated searching is sufficiently similar to real-life, unmediated searching, such that the latter can be analysed through the former. However, the author considered the obtrusivity of the search intermediary too problematic given the nature of the research question – it is common practice for

search intermediaries to assist searchers conceptualise, and re-conceptualise, their search goals.

This left transaction log analysis. The author examined the logging facilities of those OPACs in libraries which might have made logs available. The library OPAC in the pilot study logged only at the collective level – it was impossible to identify which terminal a query was performed on, or if it were the product of a remote searcher. Nor was there any output logged. Such drawbacks are typical of OPAC loggers. Logs produced by library and other information systems often fail to demarcate search sessions. Even with the assistance of interviews before or after each session, exact demarcation can sometimes still prove difficult, and manual timings are of no use if a log fails to record each query's exact time of input.

The rudimentary nature of the transaction logs produced by the library management systems to which the author had internal access, was therefore a major problem. They could not even be reliably used in conjunction with other methods, such as interviews, due to their lack of session demarcation. Further, they failed to supply output data. Other methods such as interviews and protocol analysis did not fully supply this information either, yet this was central to an examination of users' response to different system output in terms of search redefinition.

Although it became clear to the author that the transaction logs of the IR systems initially targeted for study were too crude, there were still logs of other types of IR systems to consider, if access could be obtained. These systems were not hosted by the libraries, but were offered to patrons as remote-access reference services. Such services might be subscribed to by tens of thousands of libraries around the world, most accessible over the Internet. The author was able to gain access to logs of one of the largest of these, namely, OCLC FirstSearch.

While some authorizations to the FirstSearch service are IP-based, and as such can be traced to a particular PC in a library, the author did not have access to, or knowledge of, a particular PC with its own FirstSearch authorization and which

was dedicated to FirstSearch users. This meant that analysis of the FirstSearch logs could not be complemented by other methodologies, such as interviews or protocol analysis. If the author were to rely on transaction analysis alone, for the core part of this research, then a log with copious amounts of data would be preferred, to offset the lack of triangulation through complementary methodologies. The FirstSearch log did provide substantial amounts of data.

The FirstSearch logs showed more promise than the OPAC logs in a number of respects. First, search sessions were demarcated through the use of authorizations – the log grouped together queries based on the same authorization. The queries were logged in chronological order, with exact times of input (that is, when they reached the server). Second, while the log did not include any details of output, except for numbers of hits, there was a means by which the database on which the user was searching could be identified, at least in some cases, and this could be used to determine different types of output at a broad level. Third, FirstSearch is used by a wide range of end-users, not limited to any particular kind of library, nor country or culture, nor to any particular discipline or profession, nor to any level of searching expertise. Fourth, the service includes access (potentially) to both full-text and bibliographic databases, typifying modern information provision.

A transaction log analysis based on the FirstSearch data was considered a way in which the research question could be answered. Although it might not be the ideal way, the alternatives did not appeal to the author for the reasons related above. Interview surveys or protocol analysis would be based on less data, which would not necessarily be more valid. They could not be readily supplemented by logs due to technological limitations. Mediated search studies were not based on the typical end-user searching performed on so many IR systems today. Technological limitations also ruled out online questionnaires.

A transaction log may not present us with a full picture of the searcher's thought processes, but it is assumed in this study that there is a significant correlation between queries input into an IR system, and thoughts about that IR system and its output. This study will go further than many TLAs by linking queries with

information goals, instead of strategy. That is, this TLA will investigate what a searcher is using a system to find, instead of how they may be finding it (or not finding it). This may require more interpretation of the log, but perhaps no more so than, say, content analysis. Ultimately, all analyses of thought processes require interpretation of the data derived from those thought processes, be they transcriptions of a protocol or a discourse or a set of responses to an interview, or written responses to online questions. All language is one step removed from the thoughts which produce it.

Further discussion of the issues that have been raised by TLA is included in the following review of IR research that has made use of this methodology.

5.2 Review of TLA Research

Notable reviews of studies based on transaction log analysis have been carried out by, amongst others, Peters (1993), Wyly (1996) and Blecic et al. (1999). TLA has been utilized since the 1960s, initially to monitor systems and since the late 1970s, to examine the ways systems are used. As systems began producing more descriptive and larger logs, so TLA became a popular research methodology in IR, although its credentials have sometimes been questioned (Peters 1993). The logs of OPACs and other IR systems have been studied to shed light on what users were actually doing, as well as on how experimental systems fared. In particular, transaction log analysis has been used by librarians to investigate ways to reduce search failure, either in terms of user education or in terms of catalogue design (or redesign).

Peters (1993) characterized TLA as still in the "frontier era." Many of the earlier TLAs focused on the basic elements of transaction: the frequencies of various search types, the character of the search terms, and the number of hits obtained. The units of analysis often lacked sophistication. "Unsuccessful searches," for example, have often been defined in terms of zero hits (even by Peters 1989). Thus Ballard and Smith (1992) found zero hits resulting from 25% of searches on their OPAC at Adelphi University. Also relevant to this study is the finding by Martin, Wyman and Madlock (1983) that 25% of title keyword searches on

their OPAC were, in fact, subject searches. The preponderance of keyword searching, as opposed to searching on subject indexes, is clearly evident in more recent OPAC use (Hildreth 1997).

Even Jansen and Spink (2000), in their study of the Excite logs, answered fairly straightforward questions: searchers appear to examine a small number of the documents they retrieve, with most searchers looking at a maximum of 10 items; Boolean usage is about 8 per cent. They define search failure rate as deviation from the published syntactic rules of the IR system. Jansen, Spink and Saracevic (2000) examined search *sessions,* but again in fairly rudimentary ways. They established the number of modifications per session, but analysed these modifications only with respect to numbers of terms; they categorized the most common query terms, but failed to interpret these terms in the context of the entire query and session – and terms here meant single words (adjacent sequences of characters), not phrases.

A lack of depth in many TLAs can be attributed in part to a lack of depth in the logs. Many OPACs do not allow for both search input and output to be recorded (Flaherty 1993). This problem has become all the more difficult to address with the introduction of the modern client-server systems (Jones, Gatford & Do 1997). Indeed, some modern WebPACs offer fewer logging capabilities than their predecessors (Blecic et al. 1999:529). Sophisticated, commercially available front-end utilities are rare – Flaherty (1993) identified only one such recorder, "Total Recall," and this covered only search input.

Few TLAs have examined the contents of whole search sessions chronologically. One reason for this has been the summary nature of many logs, and the difficulty of demarcating sessions between users (and terminals in some cases). Individual users cannot be identified in the log, because they do not log in and are followed by another user at the same terminal. Some TLAs have been able to solve this problem by timing users' sessions at the terminal and matching these times with those on the logs (where provided), or by asking users about the subject(s) of their session and then identifying those queries relating to this subject(s). Modern servers often assign an anonymous user code for remote access, and this was

what was used to demarcate sessions in this study. Jansen, Spink and Saracevic (2000) were also able to utilize this code, although with the Excite engine it is unclear how its continuity was identified (in the case of FirstSearch, sessions were logged on and timed out, where necessary).

A lack of search session contextualisation is typified by the study described by Blecic, Bangalore and Dorsch (1998) and Blecic et al. (1999), who found that adjustments in the OPAC interface produced greater "search success" in the short term, though this did not appear to be sustained in the longer term. The authors based their definition of "search success" on individual queries – no consideration was given to searches as series of queries. Another example of the synchronic treatment of the search, despite session demarcation, is the analysis of users' persistence in scanning search results by Wiberley, Daugherty and Danowski (1995). They examined only the most persistent search (query) per session and they ignored possible dependencies between searches within a session. That is, they failed to consider whether persistence (or lack of it) might be due in part to an evolution of search goal. (The same comment can be applied to the "quitting" study described by Wallace 1993.)

Relevance has sometimes been identified – by Wyly (1996), for example – according to the selection of full records, which provide holdings and availability information. The way in which relevance criteria "travelled" as selection progressed was an issue considered in the pilot study (see chapter 4), though Wyly's study simply looks at relevance in terms of access points that account for positive judgements.

Even when chronology has been studied, it has mostly been done from the perspective of strategy, and a static information goal is still assumed. For example, Walker and Jones (1987) examined "search state transition," where transitions are considered to be due to "negative feedback" such as results containing no hits. Likewise, Cherry (1992) examines users' reactions to and "conversions" from zero-hit OPAC results through a combination of TLA and interview methods; she assumes that such conversions are based on the original search goal. Jansen and Spink (2000) start with three basic categories of "query":

(a) the initial query (by a particular searcher); (b) a repeat query, which is the same as the initial query by the same searcher; (c) a modified query by the same searcher that is not a repeat of the initial query. Yet they do not discuss the possible reasons for the third type of query.

This omission leads us to one of the more fundamental issues concerning TLA, as recognized by Peters (1993:54), which is that logged data does not necessarily reflect the "needs, thoughts, opinions, goals, emotions and evaluations of the users." Thus it is sometimes recommended (for example, by Dalrymple 1990) that the method should be combined with others, such as the interview or questionnaire, where questions are directed at the user to elicit these needs, thoughts and so on, either by a human or through the computer screen. However, Peters (1993) does not find many instances of successful combinations of these methods. Moreover, although the data used by TLA does not include the thoughts of users directly, neither does the data that can be gained from an interview or questionnaire, as has been pointed out in the previous section. Just as logs are a product of the user's interaction with systems, so are answers to questionnaires a product of the user's interaction with a research instrument. It may be that the questionnaire is able to elicit more relevant data, but that does not mean that the log is unable to elicit any relevant data.

Ideally, data from logs should be supplemented with other data, but this does not detract from the value of TLA. Perhaps just as important as obtaining data from interviews and the like, is developing more powerful loggers that produce a more detailed record of user-system interaction. Using a screen logger, instead of a transaction logger, might be a step in this direction.

Despite its shortcomings, TLA remains a commonly-used methodology in IR, and is now increasingly being carried out on other types of IR system, apart from OPACs, such as Web search engines, which often produce very extensive logs. Website transaction loggers also offer an interesting new window into online information seeking behaviour.

TLA has been used successfully to indicate users' search strategies, so there seems little reason why it cannot be used to indicate search goals; in the final analysis, both strategies and goals are a product of users' minds. As Wallace (1993:240) observes, TLA has become an established methodology "as researchers have recognized [its] potential as a tool for delving into searchers' minds."

5.3 Interpretation of Logs

5.3.1 Introduction

As mentioned above, TLA has generally been used to answer basic questions about system or user performance – more descriptive than analytic. However, in order to answer the research question under investigation, this TLA needs to do more than summarise the logs, it needs to *interpret* them. To identify changes in information goal, we need to interpret the information goal behind each logged query. As such, each query needs to be coded according to a consistent scheme of interpretation.

Coding of transaction logs has rarely been attempted. Yet coding routinely takes places in discourse analysis and protocol analysis. For example, we have seen (in section 3.8) how Robins coded dialogue between search intermediary and information seeker to indicate modifications of a subject search; and how Dalrymple coded query reformulations in her protocol analysis.

There appears to be no theoretical reason why transaction logs cannot also be coded. It might be difficult to interpret search queries, but that does not make it an invalid endeavour. Nor would it necessarily be unreliable, if the same interpreter coded all queries according to the same coding scheme. If content analysis, protocol analysis and dialogue analysis are all legitimate methodologies, so to is an interpretive form of transaction log analysis, in the author's view. However, we first need to consider the assumptions we are making when we code queries as representations of a user's thoughts.

5.3.2 Interpreting Intention

When we discuss below the nature of various query elements, we are assuming that these expressions, input as written language into an IR system, reflect searchers' *intended* searches. Furthermore, the searcher's intended search is assumed to represent their search *goal.*

However, these assumptions are by no means certain. A user may, for example, use Boolean operators incorrectly, or misuse a term with incorrect knowledge of its meaning, or be unaware that a query defaults to a general keyword search, instead of, say, a title search. When another person interprets the logged query, especially where they possess little or no contextual knowledge of the searcher, in many cases it would be almost impossible for the users' intensions to be identified, when different from how the system will handle the query.

Indeed, the coder of the log might be unable to attribute the intended meaning of a search tem even if the user enters the query "correctly" (following the input rules of the system). The coder may attribute simply a *different* meaning to a search term from that intended by the searcher (or a coder may themselves mis-define a search term).

Likewise, searchers may *deliberately* elect not to express their search goal *exactly,* at least not initially. The reason for this could be strategic: they might wish to see what kind (and quantity) of results is obtained from a partial expression of their search goal – it might not be necessary or worthwhile to spend the additional time and energy expressing the full search goal.

While these assumptions are likely to be false on occasion, we must assume that they will be true in most cases and that in terms of when and how they are not, this is reasonably uniform across the FirstSearch databases, content differences between which will be the subject of analysis. For example, we assume that misinterpretation on the part of the coder occurs with respect to only a few searches on each FirstSearch database under investigation, and that this might result in a few cases of false identifications of redefinition (as we are defining it)

via each database and a few cases of non-identifications of redefinition via each database, but that the quantities would not be proportionately different, or not significantly so, across databases. We must assume this as we have no other evidence available concerning the FirstSearch users' goals or intended searches.

5.3.3 First-Order Categories of Query

The coding of transaction logs requires careful planning. The first question to address is, what does the coding need to show? A search session consists of one or more *initial* queries reflecting a specific search goal, each one of which may be followed up by one or more subsequent queries repeating or modifying an initial query; modifications may be strategic, or due to revision (redefinition) of search goal, or a combination of both. This study focuses on modifications of initial queries due, at least in part, to a revision of search goal – search (goal) redefinition.

5.3.4 Elements of Subject Search Queries

More specifically, this analysis focuses on *subject search* redefinition. Known-item and other non-subject searches may also lead to redefinition, but this is not to be studied here. Goal redefinition was considered most likely to occur during subject searching – which includes searches with non-topical components.

Given that search goal redefinition may occur in terms of a revision of a topical or non-topical component, we can divide the interpretation of whether a query has been modified due to strategy or redefinition into that concerning topical modification and that concerning non-topical modification. It is considered here that identification of redefinition of the topical aspect of the information goal is more feasible, and the refined index of search redefinition to be constructed in this study will be based solely on such identification. We will address this type of interpretation first, which means answering the question, When is a modification, as expressed in a logged query, likely to be due to a new topical objective, and when is it more likely to be due to a strategic consideration?

This question may not always be possible to answer, but if we analyse the topical terms in the query, we may be able to determine semantic relationships between those in an initial query and those in subsequent queries, and this may provide valuable clues.

The subject classification of search queries has rarely been attempted. One instance of subject coding of queries recorded in transaction logs is described by Pu, Chuang and Yang (2002), who created an algorithm for coding the subject content of Web search engine logs, based on the content of Web pages retrieved by the queries. The most common co-occurring terms (seed terms) were manually classified into a few broad categories, which served as the classification scheme to represent the queries. Pu, Chuang and Yang found that the automated part of the process was reasonably reliable, according to the results of a manually classified sample. While this classification can give a good indication of the broader picture of searchers' subject interests, a more sensitive classification of queries may not be so readily implemented through automation. Another instance to be noted is Hillman's coding of subject queries using the library's classification scheme (Hillman 1968). They were shown to fit the distribution of the library's collection (according to the same scheme) quite well.

When we compare one topical element with another, we shall determine whether they are synonyms or near-synonyms, or whether they are represent related concepts (which include broader and narrower ones), or whether they represent non-related concepts. If they are synonymous, or near-synonymous, then we will assume that the new term is a strategic modification. If they represent related concepts, then we will assume that some search goal redefinition is likely to have taken place – even though there are likely to be cases when a related term is entered for strategic reasons. If there is no semantic relationship between the two terms (according to the coders' viewpoint), then neither types of modification will be deemed to have taken place; rather, the new term will be considered a new initial query term, that is, the product of another search goal entirely, rather than a revised search goal.

Of course, there is bound to be an element of subjectivity involved in such classification. When is a term no longer semantically related to another term? The answer to this question depends to some extent upon the point of view of the coder. Is term *A* part of the same topic area as term *B?* This might depend on the level at which the coder defines "same topic area." Kangaroos and salmon have nothing to do with each other if the former is about mammals and the latter is about fish, but they both fall under the discipline of zoology. Perhaps the user is looking at a particular biological process common to both kangaroos and salmon? The coder must therefore take a "general knowledge" position for all their coding, rather than an "expert" position in areas with which they are more knowledgeable. The coder is to rephrase the basic question: are all the subject terms in the query, *likely* to have nothing to do with the subject terms in the initial query *as far as the searcher is concerned.* So, is the searcher likely to be making a connection between kangaroos and salmon? Or is it much more likely that this is a new search based on a separate research interest? If the latter, then we have a new initial search query, not a modified one.

The author's substantial experience in classifying and subject indexing documents across the disciplines might help in making the above judgement and in taking a reasonably objective position. Nevertheless, the element of subjectivity cannot be ignored, and some *parallel coding* was carried out in order to examine whether this subjectivity might have a significant impact on the analysis (see section 5.12).

The classification of non-topical elements of queries is even more problematic. A language qualifier "French" is exchanged for a qualifier "German." One supposes these are conceptually related, and the other elements of the query indicate that the same search goal is behind the second query, but has it been modified? In many cases, it is likely to be a strategic modification, but in some cases, the searcher may be exploring other possibilities based on a reconsideration of their information goal. More difficult still, may be the introduction of a new type of qualifier, such as a format qualifier. Again, is this strategic, to reduce the number of hits, for example, or is this because of a realization that this is the most useful format? We will assume for the purposes

of constructing a "raw index" of search redefinition here, that all new non-topical elements indicate goal modification, with only a few exceptions, such as when a location for the material is specified. In the "refined index," however, we will ignore all new non-topical elements, due to their questionable nature.

5.3.5 Coding of Queries

The coder needs to examine each element of the search query, but must interpret the conceptual relationship between queries at the level of the query. That is, the query elements must be considered as a whole, when determining a query's relationship with a previous query, as it will be assumed that each query is intended to express one unified search goal, and not several all at once. Thus each query on the log will be assigned a code. This code will be based on its conceptual relationship with an initial query, unless the query is itself deemed to be an initial query, or is deemed to be a known-item query (see section 5.10 below).

It is important to bear in mind that all query modifications are to be defined in terms of an initial search, which may not necessarily be the immediately previous query. The extent of redefinition, which these modifications are supposed to reflect (to some extent), is defined here in terms of the number of *new* definitions of information goal that evolve following interaction with the IR system. Hypothetically, a searcher may move back and forth between two definitions and in a sense this would make for many "redefinitions," but still only two definitions. This analysis is concerned with *subject travel* in terms of how many points are visited, not in terms of overall distance.

If a query is deemed to be related to an initial query, then a range of codes have been created to distinguish between certain kinds of modification (as expressed). This range is based on two aspects of search modification: whether modification involves topical or non-topical elements, or both; and whether the modification involves the introduction of new conceptual elements, or the elimination of original elements, or both.

These two aspects of query modification are thought to be worth identifying because they relate to the extent to which modification of expression is related to actual search goal redefinition. We have already noted that query modification – even as we have defined it for coding – probably does not always reflect redefinition; it may sometimes reflect strategic considerations. While we are to assume that the extent to which modification reflects redefinition is reasonably uniform across FirstSearch databases, it may be that modification due to strategic considerations is very common, in which cases differences in amounts of modification due to redefinition across databases may be difficult to detect. In other words, we might wish to pinpoint certain types of query modification that are more likely to reflect redefinition and compare these across databases, in order to detect any differences more readily. In this way, we would *refine* our index of redefinition, producing a "sharper" one.

Distinguishing between topical and non-topical terms presupposes that we can group query elements into *terms.* This is not always so easy. The same two words may sometimes be expressed as representations of two concepts, yet at other times as representing one concept as a phrase. The coder will be allowed to consider the context of the search session as a whole, and if still unsure as to the likely intended meaning of the user, will base his judgement on a rule (assume a phrase for title keywords; assume individual words for other keywords).

Once the terms of a query have been established, the coder needs to distinguish topical from non-topical terms. It was stated earlier that this analysis would encompass only subject searching, but that this would include searching involving (though not exclusively) non-subject components. This may be reasonably straightforward when labels are attached to terms, such as those for subject keyword or format qualifier. However, if the terms are entered as general keywords, it may sometimes be less obvious; in some cases, non-subject terms may even be correctly searched on subject indexes. For example, a knowledgeable searcher may use the term, "dictionaries," in a subject search on the WorldCat database intending to find, not records for documents *about* dictionaries, but records for dictionaries themselves. Again, guidance for the coder needs to be issued. In this case, the coder is advised to assume that a term

is a subject term, unless it is clear to him that it is likely to be intended as a non-subject term.

When determining whether a query modification involves the introduction of new conceptual elements, or the elimination of original elements, or both, the coder needs to distinguish between synonyms or near-synonyms, and other new elements. The coding will be conservative in the identification of near-synonyms: the coder must consider it unlikely that a new subject term is intended to have *any* different meaning from the original term; and non-topical terms will be considered not synonymous as a matter of course. When identified, synonyms and near-synonyms will not be counted as new conceptual elements; all other new terms of a query will be counted (except for one or two non-topical terms such as location of materials).

This aspect of query modification involving the introduction and elimination of elements will be transfigured into a variable based on the concept of "broadening" and "narrowing" of searches. That is, when a query modification involves only the elimination of terms, it will be classified as a "broader" search; when a query modification involves only the introduction of new terms (that count as new terms), it will be classified as a "narrower" search; and when a query modification involves both the elimination and introduction of terms, it will be classified as a "mixed" search.

The two dimensions of query modification described above will be combined into a single code for each modified query, since the way in which they combine provides a more accurate picture of the type of modification. It may well be significant if a search is narrower because of an additional non-subject term, or additional subject term, or both. If the addition is a non-subject term, this could well be for strategic purposes, or at least, more likely to be for strategic purposes than if it were a subject term. If both subject and non-subject terms are involved in a reformulation, then the modification's description becomes more complex. Are the subject terms those that have been added, or those that have been deleted, or both; similarly, for the non-subject terms? In this study, combinations of these two dimensions will be defined in terms of three sub-codes for subject/non-

subject/both and three sub-codes for broader/ narrower/mixed, producing nine codes altogether. The nine codes do not fully describe all the possible complexities of combining the two aspects described above, but they do identify several different kinds of modification which might be useful when investigating whether a sharper index of redefinition might be constructed.

There may be times when a query includes a term which represents a concept completely unrelated to any of those concepts represented in an initial query, within the same search session (according to the log). This query, however, will still be deemed a modification of an initial query if it also contains a topical term which *is* related or identical to a topical term contained in an initial query. Only when the query contains *only* unrelated terms will it be deemed conceptually unrelated and thus a new initial query. So the critical question for the coder to answer is: Do any of the subject terms in the query have anything to do with any of the subject terms in the initial search? If this is difficult to determine, the coder may also bear in mind the context of other searches in the same session. We shall assume that the search is *not* a new search if the coder considers there to be *any* possible semantic relationship, even if partial.

The reasoning behind this conservative rule for new initial queries is that evolving searches can travel large semantic distances, and that most (though by no means all) search sessions are performed with only one particular information goal in mind. In any case, the important point here is that the coding should be consistent and that the *same policy* is applied across FirstSearch databases. In this way, we assume that coding produces similarly small numbers of incorrect codes across databases, and that these will not significantly affect comparisons.

Further details of how the codes for different types of query were defined for this analysis, and details of the coding policy developed to support consistent interpretation, is provided in section 5.10 below.

5.4 The FirstSearch System

5.4.1 Introduction

FirstSearch is the primary reference service of the Online Computer Library Center (OCLC), the well-known library service based in Dublin, Ohio. OCLC is registered as a not-for-profit organization and promotes itself as the world's largest library consortium (in terms of member libraries). Although it first made its name building up the world's largest bibliographic database (WorldCat) through its cooperative cataloguing programme, OCLC is also a leading player in the reference database market. Many of the databases it has procured the rights to, for its customers, are also offered by other major information providers, such as Dialog, and most of these databases are available on OCLC's FirstSearch service.

There are in fact over 70 databases which can be accessed through the FirstSearch service, including WorldCat, the OCLC Union Catalogue. Most contain records of journal articles and other documents, some of which include links to the full text (these will be displayed if the user's authorization allows). Records may or may not contain abstracts, descriptors, etc., depending on their source. Apart from OCLC, major database providers represented on FirstSearch include H.W. Wilson, ProQuest and the Gale Group.

FirstSearch is usually accessed by users through the World Wide Web. Some of the databases are also offered through a Z39.50 server, but these are not reflected on the log provided to the author (fortunately, since the searching and displays would be localised). The service is available only to patrons of libraries; only libraries may subscribe to the service, on behalf of their patrons. Libraries subscribe to certain databases, rarely all of them. OCLC offers a number of "packages" covering groups of databases, with or without full-text; alternatively, or additionally, libraries may subscribe to individual FirstSearch databases. OCLC offers blanket subscriptions, whereby the library and its patrons may perform unlimited numbers of searches for the duration of the (annual) subscription period and may, for a higher fee, download unlimited numbers of

free-text documents. Alternatively, libraries may elect to pay on a per-search and per-download basis.

Although the author was initially focusing on bibliographic databases without full text, FirstSearch presented an opportunity to examine searching on both bibliographic and full-text databases; indeed, the majority of databases on FirstSearch do offer full text, providing the library has subscribed to it.

5.4.2 Searching on FirstSearch

The FirstSearch service offers three search interfaces for "basic," "advanced" and "expert" levels. The *basic* search interface presents the user with a single input box and assumes that they will perform, and intend to perform, keyword searches. Some limiting of searches is also offered after the results set is displayed, via a "limiter" icon – when the user clicks on the icon, they are given various limiting options, depending on the database.

The *advanced* search interface presents several boxes in a "form" format, the main input boxes accompanied by drop-down menus consisting of the field options, Boolean operator menus (and/or/not defaulted to "and"), limiting input boxes such as by year, "on-off" check boxes for limiting to the patron's library only (in terms of source journal holdings), and to full-text provision, and an input box to limit to libraries (holding the source journal) of particular OCLC codes.

The *expert* search interface presents just one main input box, but a large one that allows for the expression of complex queries using command language. The full command language for a typical database (Medline) is given in Appendix A. The same limiting boxes as in the advanced search interface (varying according to database) are also displayed.

Users may search on up to three databases simultaneously, specifying which ones by checking on the list of databases, which is limited according to their library's subscription arrangements (those not subscribed to are not listed); there is no de-duplicating of results, which are returned in separate sets, displayed one at a time

– the user clicks on a tab to display another set. The user may receive the results ranked by "relevance" (determined by an unspecified algorithm), or reverse publication date or reserve accession date ("no ranking"), according to their selection – reserve accession date being the default. Once a result set is received, it may be re-sorted according to various options (if more than 2 and less than 500 hits altogether): by author, date, number of libraries (holding the source journal), source, or title (and combinations thereof). There is another option to re-set the maximum number of brief records per screen, although probably most users would remain with the initial default, which is ten.

After the brief records are displayed, FirstSearch users may change to a "detailed" (full) record display and proceed through the records one at a time, or they may (more likely) click on the title of any particular citation to display its full record. They may also "mark" one or more of the records and retrieve these at a later stage (during the session), for group emailing or printing; or email or print records (in either brief or full format) individually. There are also options to display the session's search history, and to re-select databases.

If the full text is available, then the icon indicating this fact is displayed by the record; the user clicks on the icon in order to download the full text.

The fact that search results are presented in a uniform way across databases was of particular interest to the author, for it meant that there was scope for an investigation into the relationship (if any) between metadata content and the extent of search goal redefinition, given that there were some basic differences between the types of metadata that different databases presented to the user, at least at the full record level.

5.4.3 The FirstSearch Logs and their Management

Initially OCLC was to supply at least a month's worth of the FirstSearch transaction log. Ultimately, the author received five logs each representing a particular day of transactions on FirstSearch: 24, 28, 29, 30, 31 March 2002. (Additional logs apparently proved too time-consuming for OCLC to prepare.)

Even the five days of log, however, comprised 1,384,107 lines in total, recording 1,384,107 searches in 246,434 sessions.

The logs offered to the author by OCLC reflect *input* only, that is, they record the queries submitted by the users to the FirstSearch server, and also commands to the server requesting the downloading of full text. They do not, unfortunately, provide output data beyond the number of hits resulting from each query on each database. The queries were listed by search session and then by chronological order, so a history of each search session could be studied, at least in terms of input. A sample of the log received by the author is shown below.

Session	*Database*	*Hits*	*Search*
144517	WorldCat	212	(kw: community and kw: service) and kw: art
144517	WorldCat	1	(kw: community and kw: service) and kw: art and dt= "art"
144517	WorldCat	44	(kw: community and kw: service) and kw: art and dt= "url"
144518	other	2	(kw: lunar and kw: reproduction) and kw: fish
144518	other	7	kw: lunar and kw: reproduction
144518	other	88	kw: moon and kw: phases
144522	WorldCat	25	kw: Nursing and kw: Forum and dt= "ser"
144522	WorldCat	1	kw: Psycho and kw: oncology and dt= "ser"
144522	WorldCat	0	kw: psyco and kw: oncology and dt= "ser"
144522	WorldCat	0	kw: psyco-oncology and dt= "ser" ((de: australian and de: aborigines) and kw: documentary and dt= "vis") not (su= "Australian aborigines." and dt= "vis")
144523	WorldCat	6078	(su= "Australian aborigines.") NOT ((de: australian and de: aborigines) and kw: documentary)
144523	WorldCat	575	(su= "Australian aborigines.") NOT ((de: australian and de: aborigines) and kw: documentary) and dt= "vis"
144523	WorldCat	6124	su= "Australian aborigines."
144523	WorldCat	615	su= "Australian aborigines." and dt= "vis"

FIGURE 5.1 EXTRACT FROM FIRSTSEARCH LOG

Search sessions can be identified by the session code. Although in some cases users might have experienced a break in connection and chosen to log on again immediately afterwards, this would have been logged as a new session, and is treated as such in this study, since one cannot tell for sure if a session

chronologically close was performed by the same user, even if the search were exactly the same (it could be a fellow student with the same assignment, or perhaps a class practising searching on FirstSearch).

The five logs representing the five days were so large that they could not be combined on a single Microsoft Excel spreadsheet; in fact, they each had to be further divided up, into between three and eight sections. It was necessary to do this as parsing and other operations could only be performed on a single spreadsheet in Excel.

The reason why Excel was chosen is threefold. First, it is a tool with which the author is familiar and which is readily available to him. Second, and more importantly, the data on the log needed to be maintained in its search-chronological order for the purposes of manual coding. This meant that spreadsheet software was preferred over a database package, such as Microsoft Access. Third, the competing commercial spreadsheet packages, besides being unfamiliar to the author, do not accommodate any more rows per sheet than does Excel – the maximum for Excel being 65,536 rows.

The author ended up with 24 sheets of data from the logs. Since some of the sheets were partially filled, there were an equivalent of approximately 21 full sheets altogether.

5.5 Investigating Redefinition across FirstSearch Databases

In order to analyse the impact of output on search redefinition, differences of output need to be identified. Although there is no indication in the log of what the searcher was presented with, there was a way in which the database on which the queries were made could, in some cases, be identified. As mentioned in the section above, this could allow for predictions of what a searcher is *more likely* to see, as different databases contained records featuring different types and amounts of metadata. Since these differences are only obvious at the full record level, we must assume here that significant numbers of full records are displayed by the searchers on these different databases.

Apart from WorldCat, the FirstSearch databases were not specified in the logs provided to the author, as the database labels had been stripped by OCLC, since the organization was not at liberty (according to its lawyers) to pass on information about the searching of (and thus indirectly the performance of) specific databases, which were mostly non-OCLC owned commercial products. The label for WorldCat has been retained – this was an OCLC-owned database. Given that WorldCat represents the only "library catalogue" type of database on FirstSearch, a comparison with the other databases as a whole, most of which are indexes to articles, might of itself be meaningful, but the author wished to take the study further, if possible, comparing search behaviour on different FirstSearch databases outside of WorldCat.

Fortunately, other databases could nevertheless sometimes be identified, or at least the likely candidate could be identified, through the entries in the log which represented commands for the downloading of full-text articles. An example of an entry that shows the downloading of a full-text article is provided below. It is identifiable by the prefix "sc=".

144673 other 1 sc="1061-4303 199901 02 71 1 18 TDKOMI ?"

Each of these "sc=" lines records the downloading of a particular article, specified through its SICI (Serial Item and Contribution Identifier Standard), or SICI-like, number. In the example above, 1061-4303 refers to the source journal's ISSN, 199901 02 to the volume's chronological designation (Jan/Feb 1999) and 71 to its number, 1 18 to the pagination (1-18), and TDKOMI to the article's title (the initial letters of each word). (The question mark at the end of the quotation is an irrelevant product of the FirstSearch indexing).

These entries almost always include the ISSN (International Standard Serial Number) of the periodical in which the article was published. This provides a vital clue as to the identity of the database, since it was possible to establish a list of journals indexed on only one of the FirstSearch databases (it was not quite enough for a database to be the only one offering the full text of a periodical, since it was possible for users to still access the full text via another database on

which the periodical was indexed, if they subscribed to both databases on FirstSearch).

While one could look at differences, if any, of search behaviour exhibited by sessions involving the downloading of full text versus those that do not, it is very possible that there would be other more important factors at work: a lack of downloading of full-text may be because of a lack of relevant results; it may be because users are not interested in full-text and as such may approach their searching differently from those who are; or it may be that users are not subscribed to full-text options, and as such, again, may approach their searching differently from those who are (the logs do not indicate whether or not the searcher's library has subscribed to the full-text option on a database, so this variable cannot be readily controlled). On the average, those with full-text access might perform more searching, for instance, compared with those who will need to follow up their searching on FirstSearch with a search for the physical copies of the relevant journals in their library.

We can in fact reduce the impact of different search behaviour caused by bibliographic versus full-text objectives (on the part of users), by examining only those search sessions where there is no downloading, or only those sessions where there is downloading. At the same time, we can reasonably assume that if differences in output encountered by the searcher do indeed influence the amount of search redefinition, that such influence is primarily due to differences in metadata content, rather than differences in full text, as searchers do not normally spend time reading full text during the search session itself.

We cannot eliminate differences in users' intentions altogether, of course. Many users with access to full text, will nevertheless wish to make a note of references to documents that FirstSearch does not make available to them, but which they may find elsewhere (such as in their physical library). On the other hand, some users (particularly those working from home, perhaps) may be interested only in those documents they can download and use immediately. Again, we need to assume that this particular aspect of user variability is reasonably constant across databases.

This analysis concentrates on those sessions in which the user downloads one or more full-text article, as specific databases can only be identified in such cases – and only in a proportion of such cases at that.

While specific databases could be identified for queries recorded immediately prior to the downloading of uniquely indexed articles (on FirstSearch), it is recognised that some sessions would likely have involved more than one database. Replicating the same searches in a session on a specific database and observing numerical results close to or matching all of those shown on the log, would indicate the likelihood of a predominant, if not total, use of that particular database (although some results could apply across databases). In cases where there were significant mismatches, then it would be safer to discard such sessions from a cross-database analysis, since search behaviour might be affected by the content produced by the other databases searched. (Replicated results might be approximate instead of matching for some searches as during the interval between original search and replication, additional articles would have been indexed, and possibly others de-indexed, which might affect the results, though probably only slightly given that this interval was only a few months.)

In the event, some sessions were discarded before replication of them was attempted. In these sessions, a query was repeated exactly and consecutively on the log, and this would most likely be due to a submission to multiple databases simultaneously (given the maximum of three databases for simultaneous searching).

5.6 Metadata Displayed on FirstSearch Databases

5.6.1 Beyond the Brief Record

To investigate differences of metadata content across databases, as a factor on search redefinition, a detailed survey of these differences is required. As mentioned previously, brief records, as displayed on FirstSearch, do not differ markedly across databases. In most cases, the following elements are presented to the user:-

Title
Author (if any)
Source (or publication details if book, etc.)
Libraries (number holding the source journal, etc. as registered in WorldCat)

However, different elements and amounts of metadata are displayed on the full records from different databases. Within many of the databases, records vary in terms of the elements included, but a *typical* record can often be discerned. For instance, in a database, some records may not include an abstract, while some do include identifiers, but the vast majority of records do include an abstract and do not include any identifiers. In this study, given that we do not know which full records (if any) were displayed to the user, we need to assume that over several search sessions, users (as a whole) were exposed to metadata from the full records from the various databases, the content of which *approached* the typical records from their respective databases. So for example, users searching on a database in which 90% of records contain an abstract, are assumed to encounter more abstracts than do users searching on a database in which only 10% of records contain an abstract – on the average.

Statistically, we can estimate the minimum number of search sessions where we could reasonably expect the differential between the amounts of an element encountered to be a certain degree. Extending the above example, if we took just one pair of sessions, one session on database A with 90% abstracts and one session on database B with 10% abstracts, and each session included, say, five full record displays, then if the number of records in both of the databases is reasonably large, the probability of database A displaying more abstracts than database B is approximately 1 - 0.0001469 = 0.9998531 (see table 5.1 for calculation).

In other words, with just one session on each database, we might reasonably expect more abstracts encountered on database A. However, we would want the differential to be a significant amount such that it might be considered as the cause of a difference in search behaviour, so we might require a reasonable

expectation of, say, at least twice as many abstracts displayed from database A than from database B. Still, with just one session on each database, with five full records displayed in each session, the probability of this would also be very high: 0.985345 (see table 5.2 for calculation).

Of course, to be *sure* of a differential, one would need to have a database with 100% abstracts and another with 0% abstracts; we then simply need a mean number of full record displays for each database of more than zero.

We need to assume that the number of full record displays is enough to bring out the differences in metadata content between databases, and also sufficient to produce significant differences across databases, that is, significant such that there is a reasonable chance for these differences to result in differences in the amount of search redefinition. If few full record displays occur, then it might be that there is still a relationship between metadata content and redefinition, but that the analysis may fail to find it. The lack of evidence would not particularly suggest that metadata content was not a significant factor. When deciding upon the sample sizes, we need to bear this point in mind – that not all search sessions include a full-record display. However, it is assumed here that the majority do.

TABLE 5.1 PROBABILITY OF DATABASE A DISPLAYING MORE ABSTRACTS THAN DATABASE B

A with abs	A without abs	B with abs	B without abs	p1	p2	p3	p4	Choose A	Choose B	P
4	1	5	0	0.6561	0.1	0.00001	1	5	1	3.2805E-06
3	2	4	1	0.729	0.01	0.0001	0.9	10	5	3.2805E-05
3	2	5	0	0.729	0.01	0.00001	1	10	1	7.29E-07
2	3	3	2	0.81	0.001	0.001	0.81	10	10	0.00006561
2	3	4	1	0.81	0.001	0.0001	0.9	10	5	3.645E-06
2	3	5	0	0.81	0.001	0.00001	1	10	1	8.1E-08
1	4	2	3	0.9	0.0001	0.01	0.729	5	10	3.2805E-05
1	4	3	2	0.9	0.0001	0.001	0.81	5	10	3.645E-06
1	4	4	1	0.9	0.0001	0.0001	0.9	5	5	2.025E-07
1	4	5	0	0.9	0.0001	0.00001	1	5	1	4.5E-09
0	5	1	4	1	0.00001	0.1	0.6561	1	5	3.2805E-06
0	5	2	3	1	0.00001	0.01	0.729	1	10	7.29E-07
0	5	3	2	1	0.00001	0.001	0.81	1	10	8.1E-08
0	5	4	1	1	0.00001	0.0001	0.9	1	5	4.5E-09
0	5	5	0	1	0.00001	0.00001	1	1	1	1E-10
										0.0001469

TABLE 5.2 PROBABILITY OF AT LEAST TWICE AS MANY ABSTRACTS DISPLAYED FROM DATABASE A THAN FROM DATABASE B

A with abs	A without abs	B with abs	B without abs	p1	p2	p3	p4	Choose A	Choose B	P
5	0	2	3	0.59049	1	0.01	0.729	1	10	0.043047
5	0	1	4	0.59049	1	0.1	0.6561	1	5	0.19371
5	0	0	5	0.59049	1	1	0.59049	1	1	0.348678
4	1	2	3	0.6561	0.1	0.01	0.729	5	10	0.023915
4	1	1	4	0.6561	0.1	0.1	0.6561	5	5	0.107617
4	1	0	5	0.6561	0.1	1	0.59049	5	1	0.19371
3	2	1	4	0.729	0.01	0.1	0.6561	10	5	0.023915
3	2	0	5	0.729	0.01	1	0.59049	10	1	0.043047
2	3	1	4	0.81	0.001	0.1	0.6561	10	5	0.002657
2	3	0	5	0.81	0.001	1	0.59049	10	1	0.004783
1	4	0	5	0.9	0.0001	1	0.59049	5	1	0.000266
5	0	2	3	0.59049	1	0.01	0.729	1	10	0.043047
5	0	1	4	0.59049	1	0.1	0.6561	1	5	0.19371
5	0	0	5	0.59049	1	1	0.59049	1	1	0.348678
4	1	2	3	0.6561	0.1	0.01	0.729	5	10	0.023915
										0.985345

5.6.2 Full Record Contents

Full-record displays consist of those fields present in the brief citations, and commonly five or ten others. However, *which* others vary considerably across databases. The task now was to identify substantial and consistent differences amongst the databases.

In addition to the brief citation fields, the metadata elements listed below are displayed on at least some of the databases, in the full records. Some database-specific metadata elements (indented in the list) are classed under their generic categories. Identifiers are defined here, as they generally are in FirstSearch, as uncontrolled subject terms. It should be noted that slightly different labels are displayed on different databases (e.g. "citation" vs "cited reference"), but it is assumed that this is unlikely to influence search behaviour.

Abstract
Article/Document type
Author affiliation (institution)
Citation
Classification/subject code
    Company number
    Industrial codes/groups
    Product classification
Conference name
Control number (e.g. accession number, Medline number)
Corporate name/author
Descriptor (major/minor/special, etc.)
    Chemical indexing
    Chemical substance
    Geographic coverage
    Marketing term phrase
    MESH heading
    Scripture citation
    Statute citation
Edition

Full-text availability

Identifier (subject)

- Company name
- Named corporation
- Named person
- Product name

Intended audience/Age group

Language

Notes

Number of references

Physical description

Reviewed journal

Series

Special feature

Standards number (e.g. ISBN, ISSN)

Table of contents

Uniform title

5.6.3 Selection of Displayed Metadata for Analysis

It was supposed that most of the above elements, when displayed, would rarely be the cause of redefinition of a search goal, especially with respect to subject searching. Subject-related format types (e.g. "obituary" for a biographic database) appear in a few databases, but this is unlikely to cause topical refinements; citations are also present in some, but only include authors and source details, not titles; classifications and "categorizations" are also sometimes present, but are much less likely to be noted than are the more specific and clearer subject descriptors.

Subtitles were often included in the full record (but not in the brief record), when they existed (on the document). A substantial lack of subtitles on the full-record displays might make a significant difference with respect to subject search behaviour, but the databases that were ultimately selected for analysis all included many subtitles.

Most of the metadata elements listed above consist of no more than one line, when most full records display at least ten or fifteen lines (those with abstracts, many more). More likely would an abstract or several lines of descriptors or identifiers catch the eye of the user. It was these three elements – abstracts, descriptors and identifiers – which was thought most likely to have a significant impact on the amount of search redefinition taking place. It was clear that the provision of these three elements varied considerably across databases, and they were chosen as the main focus of the analysis.

Two aspects of the metadata elements became the subject of analysis: presence or absence; and length. Both the frequency of abstracts or descriptors in a database and the extent of them (e.g. number of words per abstract) might make a difference with respect to the amount of redefinition.

5.7 Control of Other Variables

5.7.1 Quality of Metadata

As well as the quantity of abstracts, descriptors and identifiers, the quality of this metadata, and other types of metadata, such as titles and subtitles, might also lead to more or less redefinition. It may be that two different abstracts of equal length on the same article, might cause redefinition to significantly different extents. The *specificity* and, to some extent, *style* of abstracts and titles might vary according to *subject.* Similarly for identifiers, which are commonly taken from titles/subtitles and/or abstracts, or employed as a kind of short-hand abstract. For discussion of the "subject factor," see section 5.7.6 below.

While abstracts, titles/subtitles and identifiers are written in natural language, the controlled vocabulary of descriptors may vary across databases in ways beyond that of subject. Some vocabularies are larger than others, some are formatted in a pre-coordinated style, others in a single-term post-coordinated style. However, the nature of the descriptors presented to the FirstSearch user is not solely related to a particular vocabulary, but also to the way in which they are used by the indexers. More exhaustive indexing could result in more descriptors despite a smaller vocabulary. Examining the nature of the indexing for each database would be a very steep task. Instead, it will be

assumed that this variable would not make a large impact on redefinition, in comparison with the many other possible factors, such as presence or absence of an abstract. However, those databases with unusual subject indexing (either very extensive, with several lines of major and minor descriptors, or very limited, with never more than one descriptor in each full record) will be excluded from the analysis.

We shall assume that if other aspects of metadata's content influence the extent of redefinition, it is reasonably uniform across databases.

5.7.2 Other Metadata Elements

The presence or absence of other types of metadata might also be redefinition factors. However, the other metadata elements were found to appear across databases representing the various combinations of abstracts/descriptors/identifiers, such that even if they were to have a slight effect on redefinition, this would likely have little bearing on the analysis.

Most metadata elements did not vary significantly in terms of their extent – they were usually less than one line in length. An exception was title length, which could conceivably affect subject redefinition. Fortunately, while title lengths were found to vary a little across databases, including those that were finally selected for analysis, the variation was not found to correlate with the presence or absence of abstracts, descriptors or identifiers, so that any significant correlation between presence and/or extent of abstracts, etc. and redefinition should be independent of title length.

5.7.3 Other Presentational Aspects of the Metadata

Metadata elements, when they are present, are generally displayed on FirstSearch in the same order down the screen, irrespective of database. It should also be noted that all the elements are displayed in the similar font, and most in sentence case.

A metadata element's influence on redefinition might be increased if it is presented as a hyperlink. In FirstSearch, descriptors and authors are presented as hyperlinks; there are also buttons that retrieve "related records" based on the same author(s) or descriptor(s).

The hyperlinking nature of the author elements are unlikely to have much impact on subject search redefinition. However, the hyperlinking nature of the descriptors may well do so. Since the descriptors are hyperlinked in every database, there is no issue of control here, but it should be noted that if the descriptors be found to make a significant difference to the amount of redefinition, a not inconsiderable part of this may well be due to their hyperlinking nature on the FirstSearch interface.

We should also note the option on FirstSearch for search results to be shown according to a relevance-ranking algorithm. This is not the default option, which is reverse accession order, but it would no doubt have been used by some searchers. Even if precision were to affect redefinition significantly (see the section below for more on this issue), we will assume that the use of the relevance-ranking algorithm occurs reasonably uniformly and is similarly effective across databases. We will also assume that the default ordering of results, and the other sorting options, would give rise to more redefinition on one database than on another.

5.7.4 Differences in Retrievability of Metadata

A major concern in terms of control is the variability of results *retrieved,* as opposed to displayed, on different databases due to differences of metadata content. For example, databases with abstracts may produce more or less recall or precision compared with databases without abstracts, and rates of recall or precision may have an effect on search redefinition. It may be hypothesised, for instance, that less precision encourages users to reflect more on their search goals, and more reflection may lead to more redefinition. Indeed, there is some evidence to support this hypothesis in the pilot study related in chapter 4. On the other hand, it is by no means clear that this would necessarily be the case. It may be that a distinct lack of success, in terms of meeting the search goal, simply heralds the end of the search session, while it is *partial* success that spurs the searcher on to "greater heights." (We may note here that Spink and Greisdorf (1997) found that partial relevance often led to shifts in information problem definition.)

In FirstSearch, subject retrieval is performed, in most cases, through a "keyword" *(kw)* search or a "subject" *(su)* search/browse. Databases do offer other indexes to search on (e.g. abstract index), but the FirstSearch logs indicate that very little in the way of such

refined searching actually takes place, except in terms of supplementary limiters. The metadata elements indexed in FirstSearch for *su* searches that are responsible for many of the hits for subject searches are the descriptors and identifiers (abstracts are not indexed in the "subject" index). In the case of *kw* searches, the most fruitful metadata are abstracts, titles, subtitles, descriptors and identifiers.

As a caveat, it is noted that there are small differences in terms of the stop-words established for the indexes of the different databases. However, these content-less words would not normally be subject searched on, so the differences, in comparison with all the words that are indexed, are most unlikely to be significant here.

Both the quality and quantity of words indexed from the abstract, descriptor and indicator fields are likely to have a bearing on precision and recall, although there may be other factors such as the kinds of searches input into different databases. Even if different levels of precision and recall are produced by different databases, it is unclear in what way, if any, this would affect search redefinition. Nevertheless, we shall hypothesise that greater retrievability of records through more index entries leads to a lack of precision which, in turn, leads to more search redefinition.

The testing of this hypothesis could be undermined by other differences between the databases' metadata and the ways they are searched on, offsetting the effect of retrievability on recall and precision, and so this investigation compares different types of search (*kw* and *su* searches) on the *same* database. The *kw* searches would retrieve on titles and abstracts as well as descriptors and identifiers, whereas the *su* searches would not. If the two types of search showed no significant difference in terms of (immediately) subsequent search redefinition, this would indicate that the retrievability of records has little, if any, effect on search redefinition. An assumption is made here that *su* searchers' behaviour is generally similar to *kw* searching behaviour, or at least not in a significantly different way with regards to redefinition.

5.7.5 Differences in Low Hit Rates

Apart from the retrievability of metadata, other factors may have an impact on what may crudely be termed "search failure," that is, low hit rates. The size of a database (in

terms of indexed words) is likely to have an influence, as is the kind of searches input by the users (for example, do they employ valid controlled vocabulary?). Hit rates, as well as levels of precision and recall, may have an impact on search redefinition, and it may be that the FirstSearch databases under analysis produce, at least in these logs, different distributions of hits.

Whether hit rates do influence search redefinition (in the terms of this study) can be investigated directly by analysing some of the searches on the logs resulting in low numbers of hits, including zero hits: are they more or less likely to result in redefinition? If there is a correlation, it is more likely to occur with lower numbers of hits, as studies have shown that many users give up examining records after one or two screens (as shown, for example, in the Excite study by Jansen, Spink and Saracevic 2000).

In this analysis, we will not attempt an exhaustive investigation as to whether *any* particular hit rate affects redefinition, but we will investigate whether certain low levels, as shown on the logs, significantly correlate with higher or lower frequencies of redefinition.

5.7.6 Subject of Databases

An important factor which may affect search behaviour is the subject of the database and corresponding subject-orientation of the searcher. Therefore, the relationship between the subjects of databases and the amount of redefinition they generated was examined. When selecting those databases to be used in the analysis of redefinition and metadata content, those that represented a wide range of subjects, or groups that represented such a range, were preferred, although in practice the choice was a limited one.

5.7.7 Other Database Differences

Another possible factor that varies across database is document currency. One database was eliminated due to its archival nature, namely the Alternative Press Index Archive. Other databases would have slightly different mean ages for their documents

(represented by their records), but they all have many records for documents recently published. The (initial) default ranking of the brief citations is by reverse accession order, which approximates to reverse chronological order. Thus most subject searches on any of these databases (except for the archival database discarded) would produce a screen or more of citations for recent documents, published in the last year or two. It is thought unlikely, therefore, that slight differences in the age distributions of each database's documents would have a significant impact on overall search behaviour.

It is also noted that the rate at which new records are added to the FirstSearch databases varies. It is conceivable that a less frequently updated database would have fewer successive searches performed on it, and this in turn might affect measures of search redefinition. However, the updating frequencies differ in terms of weeks rather than years and are probably of little consequence to most users – indeed, most users are probably unaware of the differences. Furthermore, most users will not rely on FirstSearch for their information needs, and will seek information and conduct research through a wide range of other media. In this context, one or two more successive searches on database A compared with database B is unlikely to be of much significance.

5.7.8 Types of User

Given that we are comparing search behaviour across databases, we need to consider any other factors related to the *use* of the various databases that might result in differences in this behaviour. The *type* of user may well vary according to subject matter of the database and the kinds of library that subscribes to it, and different types of users may behave differently on FirstSearch. For example, public library users are presumed less likely to use advanced searching techniques, and they are less likely to perform "comprehensive" searches, compared with university library users. Those databases primarily targeted by OCLC at non-academic audiences were thus discarded in the initial de-selection process. While some public, school and special libraries subscribe to the remaining databases, the academic library sector accounts for the vast majority of users, and thus, we presume, a very high percentage of searches on each of the databases.

In terms of academic libraries, one might wish to distinguish between libraries of universities and those of colleges of pre-university education. However, the former would predominate, it may be safely assumed, in the subscriptions to all of the databases under examination; besides, there appears little reason to suppose significantly higher numbers of pre-university college libraries subscribe to any specific database, for there is none particularly geared to further education students, while all the databases may be relevant to certain of their courses. Within the university user group, one may wish to further distinguish between undergraduates, postgraduate students, and faculty. However, again, all the databases in question are suited to all three types of user, and none can be considered as representing especially "advanced" literature.

It is unlikely that searching on particular databases would be slower, in terms of response time, than on others, and it is assumed in this study that any differences in response time would not, in any case, translate into significantly different amounts of search redefinition. It is also assumed that although some libraries may pass on more of the financial costs (telephonic and/or OCLC charges) to users, this would not lead to a differential across databases that would significantly affect search behaviour. Likewise, it shall be assumed in this study that any differences between databases in terms of the proportions of search sessions undertaken from outside of the subscribing libraries, are not significant, as long as one excludes those databases aimed primarily at public libraries or schools. It is also assumed that the five days logged were not atypical in terms of the differences in redefinition that might be detected across databases.

The subject orientation of the user may well have an effect on search redefinition and is obviously related to database. Scientists might perform more precise initial searches compared with humanities scholars. Differences in behaviour on databases for different disciplines may impact on redefinition. For treatment of the "subject factor," see section 5.7.6 above.

Another assumption that is made in this analysis is that the overall information-seeking contexts of the typical users of the different FirstSearch databases are not significantly different. It is not the case, say, that database A is always its users' first choice of IR system, whereas database B is only its users' third or fourth, last-resort choice. Nor is it the case that users of database A are usually in the early stages of their research on a

given topic, while users of database B are usually at what Kuhlthau and Cool (1992) might call the "solution presentation" stage.

One final, important assumption needs to be recognised here: the FirstSearch databases are designed for end-users, but some libraries may offer them through a mediated search service. We are assuming that the amount of mediated searching does not vary significantly across databases, given that this amount is probably quite small. A majority of FirstSearch subscribing libraries are North American academic libraries, and Tenopir's survey of such libraries conducted shortly before the data was collected for this study indicated that FirstSearch was firmly an end-user service, with no libraries using it for intermediary searching (Tenopir 2001).

5.7.9 Environment External to FirstSearch Session

Although the FirstSearch service is easily accessible for many end-users, there is another aspect of its WWW environment that should be considered: users may be able to open other Windows for other activities simultaneously. It is assumed in this study that simultaneous activities are minimal, although it is also likely that some do occur and that in some cases information or "inspiration" is encountered outside of FirstSearch that affects the direction of searching; it is further assumed, however, that such changes of direction caused by external factors are reasonably constant across searching on the FirstSearch databases.

5.8 Selection of Data for Analysis

5.8.1 Preliminary Selection of Databases

A summary of the characteristics of the FirstSearch databases were available on OCLC's FirstSearch website; the summary is currently located at: http://www.oclc.org/support/documentation/FirstSearch/databases/dblist/

Some of the databases were straightaway excluded from the analysis, due to their inapplicability, their non-standard nature, their non-academic target audience, or because they did not contain metadata or point to full text.

The remaining databases were investigated in terms of their metadata content – in particular, their descriptors, identifiers and abstracts. Over 100 sample records, derived from content-less queries (e.g. "through" as a keyword) on each database, were examined for their metadata content, while relevant details about the databases were gleaned from information available on the FirstSearch documentation, the websites of the database vendors, and the Dialog Bluesheets. Table 5.3 shows the results of this investigation.

TABLE 5.3 KEY METADATA CONTENT OF FIRSTSEARCH DATABASES

Y = most records contain the element
S = some records contain the element, but not most

Database	Abstracts	Descriptors	Identifiers
OCLC ArticleFirst			Y
Electronic Collections Online			Y
Applied Science & Technology Index		Y	
Art Index		Y	
Biological & Agricultural Index		Y	
Education Index		Y	
General Science Index		Y	
Humanities Index		Y	
Library Literature		Y	
Social Sciences Index		Y	
Biography Index		Y	Y
Business Dateline		Y	Y
MLA International Bibliography		Y	Y
Arts & Humanities Search	S		Y
Health and Wellness Information	S	Y	
AGRICOLA	S	Y	S
Contemporary Women's Issues	S	Y	S
Business & Industry	S	Y (extensive)	Y
CINAHL	S	Y	Y
PsycINFO _1887--	S	Y	Y
ABI/INFORM	Y	Y	
Applied Science & Technology Abstracts	Y	Y	
Art Abstracts	Y	Y	
BasicBIOSIS	Y	Y (extensive)	
EconLit	Y	Y	
Education Abstracts	Y	Y	
General Science Abstracts	Y	Y	

GEOBASE	Y	Y (limited)	
Humanities Abstracts	Y	Y	
PAIS International	Y	Y	
Periodical Abstracts	Y	Y	
Social Sciences Abstracts	Y	Y	
Wilson Select Plus	Y	Y	
Business & Management Practices	Y	Y	Y
ERIC	Y	Y	Y
Medline	Y	Y	Y
PsycARTICLES	Y	Y	Y
Sociological Abstracts	Y	Y	Y

Databases listed in table 5.3 were grouped according to the main metadata combinations; databases that fell outside these groups, including those with relatively high or low numbers of descriptors per record, were excluded. The groupings are listed below.

Group A (identifiers only)

Database	Abstracts	Descriptors	Identifiers
OCLC ArticleFirst			Y
Electronic Collections Online			Y

Group B (descriptors only)

Applied Science & Technology Index		Y	
Art Index		Y	
Biological & Agricultural Index		Y	
Education Index		Y	
General Science Index		Y	
Humanities Index		Y	
Library Literature		Y	
Social Sciences Index		Y	

Group C (descriptors & identifiers only)

Biography Index		Y	Y
Business Dateline		Y	Y
MLA International Bibliography		Y	Y

Group D (some abstracts, many descriptors and/or identifiers)

Arts & Humanities Search	S		Y
Health and Wellness Information	S	Y	
AGRICOLA	S	Y	S
Contemporary Women's Issues	S	Y	S
Business & Industry	S	Y	Y
CINAHL	S	Y	Y
PsycINFO _1887--	S	Y	Y

Group E (many abstracts & descriptors)

ABI/INFORM	Y	Y	
Applied Science & Technology Abstracts	Y	Y	
Art Abstracts	Y	Y	
BasicBIOSIS	Y	Y	
EconLit	Y	Y	
Education Abstracts	Y	Y	
General Science Abstracts	Y	Y	
Humanities Abstracts	Y	Y	
PAIS International	Y	Y	
Periodical Abstracts	Y	Y	
Social Sciences Abstracts	Y	Y	
Wilson Select Plus	Y	Y	

Group F (many abstracts, descriptors & identifiers)

Business & Management Practices	Y	Y	Y
ERIC	Y	Y	Y
Medline	Y	Y	Y
PsycARTICLES	Y	Y	Y
Sociological Abstracts	Y	Y	Y

The following short-list of databases was thus drawn up for collection of their logged search sessions:-

1. WorldCat
2. ABI/INFORM
3. Applied Science & Technology Abstracts
4. Applied Science & Technology Index
5. Art Abstracts
6. Art Index
7. Biological & Agricultural Index
8. Business & Management Practices
9. CINAHL
10. EconLit
11. Education Abstracts
12. Education Index
13. Electronic Collections Online
14. ERIC
15. General Science Abstracts
16. General Science Index
17. Health and Wellness Information
18. Humanities Abstracts
19. Humanities Index
20. Library Literature
21. Medline
22. OCLC ArticleFirst
23. PAIS International
24. Periodical Abstracts
25. PsycARTICLES
26. PsycINFO _1887
27. Social Sciences Abstracts
28. Social Sciences Index
29. Sociological Abstracts
30. Wilson Select Plus

5.8.2 Investigation of Data Availability

Although the data received from OCLC represented a very large number of search sessions, the majority of these search sessions did not, according to the log, include the downloading of a uniquely indexed article, and thus would not be of use for the purposes of the cross-database analysis of the influence of metadata content on redefinition.

A full sheet of the logs imported into Excel was first used as a sample in order to identify the numbers of search sessions that could be used for the particular analyses that the author wished to perform. In particular, it was necessary to estimate the numbers of sessions on each of the databases on the short-list above (section 5.8.1) that could be identified as such. It was also estimated, via sampling, that these numbers would need to be halved, following the elimination of sessions with repeated query lines, which were assumed to represent multiple-database searching (see section 5.5 above). The resulting estimations of applicable sessions, to be derived from the logs, are shown below.

TABLE 5.4 ESTIMATED NUMBERS OF APPLICABLE SESSIONS IN THE LOGS

Database	Number of sessions
WorldCat	7,304
ABI/INFORM	31
Applied Science & Technology Abstracts	0
Applied Science & Technology Index	0
Art Abstracts	0
Art Index	0
Biological & Agricultural Index	0
Business & Management Practices	31
CINAHL	0
EconLit	0
Education Abstracts	31
Education Index	0
Electronic Collections Online	0
ERIC	0
General Science Abstracts	0
General Science Index	0
Health and Wellness Information	178
Humanities Abstracts	0

Humanities Index	0
Library Literature	10
Medline	189
OCLC ArticleFirst	42
PAIS International	10
Periodical Abstracts	52
PsycARTICLES	0
PsycINFO_1887	0
Social Sciences Abstracts	0
Social Sciences Index	0
Sociological Abstracts	21
Wilson Select Plus	0

5.8.3 Sample Size Requirements

The estimated number of sessions that could be analysed was disappointingly low for many databases. The question thus arose, what would be a reasonable sample size, for any given database? This needed to be decided upon at this stage, as identifying search sessions on each of the databases involved complex filtering processes (as detailed in section 5.9 below). Obviously the sample size would need to be greater than zero, but given that the estimation was based on about 5% of the total number of sessions, there could well be databases that would end up with a handful of identifiable sessions, perhaps fewer than ten.

It should be noted here that the number of sessions does not in fact necessarily equate to the number of subject search goals, as some sessions may involve more than one initial query (based on a particular search goal), while other sessions may consist solely of known-item searches (as defined in the coding) and thus no initial queries. It is the amount of redefinition based on an initial subject query that is to be compared. However, the estimated numbers of sessions above were treated as an approximation of projected sample size.

A *reasonable* sample size is defined here as one that would make the results of significance tests *meaningful*. Results are considered meaningful if they could be acted upon in some useful way. The statistical power of a test is, in fact, not necessarily a question of "the more the better." A test may be too powerful, detecting statistically significant differences which might not be of much relevance to real life. On the other

hand, a lack of power and a result which shows no significant difference is inconclusive – it may just be that the test needs more power in order to bring out the difference.

It is of course very hard to say what might be the "right" amount of power such that the results would show some phenomena if it is "large enough" to be of interest. If redefinition occurs more frequently via database A than via database B, should it have any bearing of system design? If it is only a tiny difference, perhaps not. If the difference is large, then perhaps it should. But what if the difference is "a little"?

Furthermore, when considering what might constitute "large enough," we must remember that we are not analysing redefinition per se, but *indices* of redefinition. If there are twenty reformulations of searches in five sessions, this does *not* equate to a mean of four redefinitions per session. We are using data on query reformulation as an index – an indication – of the *relative* extents of a certain kind of search redefinition occurring through the use of FirstSearch, assuming that this index is equally representative of redefinition across databases.

Many statistical tests have formulae for power estimates, but the tests in this analysis are not being performed on *direct* data. The effect size may well be an under-representation of the actual effect of the differences (if any) found amongst the database; conversely, the "actual power" of the test may well be less than what a statistical formula might indicate. In truth, we cannot readily estimate the ideal statistical power for the significance tests we will carry out, given the indirect nature of the indices, but we shall assume that the effect size will not be grossly exaggerated because of this, and as such follow the established recommendations for sample sizes and statistical power, and accommodate higher power and greater sample sizes if they mean that other problems such as sampling issues can thus be avoided.

The exact number of samples required for a particular statistical power depends on the test, the nature of the samples, and the effect size determined – something that cannot be determined until the test has been performed. However, the recommended minimum sample size for the Kolmogorov-Smirnov 2-sample test is given by Kanji (1999) as fifteen, and Kolmogorov-Smirnov (K-S) was the main test to be used in the cross-

database analysis. Thus the minimum number of samples for each database was predetermined as fifteen.

According to this cut-off, the following databases were targeted for identification.

Database	Projected number of applicable sessions
WorldCat	7,304
Medline	189
Health and Wellness Information	178
Periodical Abstracts	52
OCLC ArticleFirst	42
ABI/INFORM	31
Business & Management Practices	31
Education Abstracts	31
Sociological Abstracts	21

TABLE 5.5 DATABASES TARGETED FOR IDENTIFICATION

5.8.4 Target Datasets

Unfortunately, not all the "metadata groups" A-F listed in section 5.8.1 are represented by the databases in table 5.5 above. Group A is represented by a generic database, OCLC ArticleFirst, while groups E and F are represented by databases in a range of disciplines: ABI/INFORM, Education Abstracts, and Periodical Abstracts for group E; and Business & Management Practices, Medline, and Sociological Abstracts for group F. Groups B and C are not represented at all, while group D is represented by only one, specialized database, namely Health & Wellness Information.

In order to facilitate more opportunity to compare of the effect (if any) of metadata combinations, datasets based on the specific databases listed above could be supplemented by datasets representing groups of databases.

Most generically, datasets could be derived from search sessions on any database except WorldCat. Two such datasets were targeted: that containing search sessions where downloading of full text occurs, and that containing search sessions where no such downloading occurs. The presence or absence of a "sc" line (and the absence of any

WorldCat labels) would make identification of sessions for these two datasets a relatively straightforward matter (sampling issues aside).

Further, search sessions could be identified which may not necessarily include the downloading of a *uniquely* indexed article, but which did include one indexed only on databases represented by one or more of the "metadata groups" A-F. The author considered that perhaps a lack of abstracts would be the most likely metadata factor on search redefinition, and so wished to address the lack of representation of groups B and C amongst the databases listed in table 5.5. The sample log was re-examined, and it was estimated that 21 sessions based on one or more group B database might be found from the five days of log. Likewise, a combined group A-C of databases, representing all those databases without abstracts, minus OCLC ArticleFirst, was investigated for its exclusive sessions, using the sample log, and the projected total was estimated as 73. Group B and this combined group A-C, which is to be henceforth called "Group Z," were thus targeted as datasets to be derived for analysis, bringing the total number of datasets to be derived for the primary analysis, to thirteen, as listed below.

1. WorldCat-only
2. Other-only without any full-text downloading
3. Other-only with full-text downloading
4. Group B databases with full-text downloading
5. Group Z databases with full-text downloading
6. OCLC ArticleFirst
7. ABI/INFORM
8. Business Management
9. Education Abstracts
10. Health & Wellness
11. Medline
12. Periodical Abstracts
13. Sociological Abstracts

Group B databases: no abstracts or descriptors

Applied Science & Technology Index

Art Index

Biological & Agricultural Index

Education Index

General Science Index

Humanities Index

Library Literature

Social Sciences Index

Group Z databases: no abstracts

Electronic Collections Online

Applied Science & Technology Index

Art Index

Biological & Agricultural Index

Education Index

General Science Index

Humanities Index

Library Literature

Social Sciences Index

Business Dateline

MLA International Bibliography

Biography Index

5.9 Production of Datasets for Analysis

5.9.1 Sampling Considerations

In the case of datasets 4-13 listed in the above section, there was not considered to be much of an issue with regard to sample size. That is, their potential sample size, if as many samples as possible were to be derived, was not considered likely to be too large (by making for too much power).

However, in the case of the first three datasets – WorldCat and non-WorldCat searches – many thousands of search sessions could be generated for coding, and this was considered both unnecessary and unbalanced. Instead, a target of about 1,000 sessions was set, such that the number of initial searches might be 500 (once the known-item

only sessions had been excluded). This was considered an acceptable number statistically.

For the creation of these three datasets, then, the question of sampling method arises. Ozmutlu, Spink and Ozmutlu (2002) identify different search behaviour and subjects at different points in time on the Excite search engine and warn that sampling techniques need to overcome this phenomenon. However, the author is of the view that search behaviour on the FirstSearch databases is likely to be more uniform over time. The three datasets were derived from different sections of time, spread across all five days: the first 50 applicable sessions recorded on the 21 Excel sheets were used (a total of 1,050 altogether). The datasets were thus considered sufficiently representative of the five day period to allow for comparison with the other datasets derived from the full five days of log.

5.9.2 Production of WorldCat Dataset

WorldCat-only sessions can be readily identified as those without any lines containing "Other" in the database column. Session code numbers with "Other" along side were identified (through a sort on the database column), then all lines with these code numbers were filtered out, leaving those sessions with only "WorldCat" in the database column.

5.9.3 Production of Other-Only Datasets

The log sheets were reduced by filtering out sessions with any "WorldCat" entries by means of the converse process described above to identify WorldCat-only sessions, thus producing "Other"-only sessions.

The remaining sessions were copied out and re-imported, and at the same time parsed so that "sc=" was separated into its own column, when found in a line. The tables were then sorted by this column and all the rows containing "sc=" copied out and re-sorted by session code number. The list of code numbers was then de-duplicated and the result added to the log sheet. Those rows which did not match any of the code numbers were filtered out, leaving those sessions that included one or more full-text downloading.

The list of code numbers was then added again to another copy of the log sheet and this time the rows which did match any of the code numbers were filtered out, leaving those sessions that did not include a downloading.

5.9.4 Production of Other Specific Database Datasets

The table of indexed journals by database was taken from the FirstSearch website in April 2002 and was assumed to reflect fairly accurately the situation during the time the sessions were logged. The table included the journals' International Standard Serial Number (ISSN) in most cases. The table was imported into Excel and sorted by ISSN and then database. Multiple ISSNs and their rows were filtered out, leaving the uniquely indexed journals represented by their ISSN, and their respective databases. The table was then re-sorted by database.

The resulting table list might not include quite all of the uniquely indexed journals, for it excludes any missing ISSNs. Furthermore, the table lists uniquely indexed journals, which is not quite the same thing as uniquely indexed articles: there may have been a few uniquely indexed articles in journals not uniquely indexed, but where only one database indexed those particular articles.

The log sheets that had been reduced to Other-only sessions were copied and re-imported into Excel, and at the same time parsed such that the ISSN following "sc=" (where it existed) was formatted in a separate column. The ISSNs of the uniquely indexed journals for each of the eight specified databases were then transferred *in turn* onto the log sheets. For each log sheet, queries containing an ISSN that matched one of the ISSNs of the uniquely indexed journals were identified, copied out, and sorted by session code number. Multiple code numbers were de-duplicated. The resulting set of session code numbers, for each database, was then added to the log sheet and all rows which did *not* match one of the code numbers were filtered out, leaving those sessions which contained the downloading of a full-text article that could be identified as having resulted from a search on a specific database.

The sessions remaining on the sheets for datasets 6-13 were then reduced further by filtering out those sessions that included an adjacent repetition of a line, which would

generally mean a repetition, or simultaneous performance, of a search on a different database.

5.9.5 Production of Group B & Group Z Datasets

The table of journals from OCLC was re-imported into Excel and sorted by database. The ISSNs for those databases in Group B were *cut* out and de-duplicated. This list was added to the remaining table of journals and those ISSNs in the list that did *not* match any of the ISSNs in the remaining rows of the table were identified and copied out: this final list thus contained the ISSNs for databases in Group B, but for no other databases.

This list was then added to each of the log sheets, after the sheets had been reduced by filtering out sessions with any "WorldCat" entries as for the processing towards datasets 6-13, and the ISSNs in their SICI numbers parsed. Rows containing an ISSN (in the ISSN column) that matched one of those in the list of ISSNs were then identified, copied out, and sorted by session code number. Multiple code numbers were de-duplicated. The set of session code numbers was then added to each of the log sheets and all rows which did *not* match one of the code numbers were filtered out, leaving those sessions which definitely included a search on a Group B database.

The same procedure was repeated for Group Z databases.

The sessions remaining on the sheets for datasets 4-5 were then reduced further by filtering out those sessions that included an adjacent repetition of a line, which would generally mean a repetition, or simultaneous performance, of a search on different databases (which could be outside of Groups B or Z).

It should be noted that in the above processes, while in some cases the lists of session code numbers generated from each log sheet could be merged, they could only be merged within the groups of sheets for each day, not across days, since otherwise some mis-identification would likely occur due to duplication of code numbers over the five days of logs – the code numbers were only unique within each particular day's log, they would be repeated in following days' logs.

5.9.6 Confirmation of Datasets

For the *refined index,* another step was taken. Once the single-database sessions had been identified, the hits on searches (if any) prior to the search before the identifying "sc" line were compared with the results produced by a simulation of the search by the researcher; likewise for searches (if any) subsequent to identifying "sc" lines. Not all searches were simulated, but at least two per session, depending on the extent of the session. Most hit sizes from the simulated searches were found to compare with the numbers logged; those sessions where a discrepancy was found, were discarded when compiling the refined index.

5.9.7 Sessions for coding

The table below shows the numbers of sessions, by dataset, which were ultimately passed for coding.

Dataset	Sample sessions for coding
WorldCat-only	1,050
Other-only without any full-text downloading	1,050
Other-only with full-text downloading	1,050
Group B databases	15
Group Z databases	61
OCLC ArticleFirst	18
ABI/INFORM	25
Business Management	43
Education Abstracts	32
Health & Wellness	186
Medline	183
Periodical Abstracts	45
Sociological Abstracts	27

TABLE 5.6 NUMBER OF SESSIONS ON THE LOGS PASSED FOR CODING

Although the Group B dataset was borderline in terms of minimum sample size (see section 5.8.3 above), it was decided that the coding could proceed in all 13 datasets, and that the analysis should be based on differences between these datasets.

5.10 Construction of the Coding System

Some of the main assumptions involved in coding transaction logs and the way in which coding might cover aspects of query modification, have already been discussed in section 5.3. More detailed description of the construction of a coding system and the surrounding policy to be implemented in the specific context of the FirstSearch log, is provided in the following section.

5.10.1 Logging of FirstSearch Queries

In order to construct the coding system, we first need to understand the way in which a query on FirstSearch is logged. There are in fact two types of search function offered on FirstSearch, at least on the advanced and expert interfaces. As well as matching on the search terms, the system may be directed to browse a particular index, that is, to display that part of the index containing the (first input) search term, or nearest to the search term, alphanumerically. When browsing, however, in order to retrieve records the user must take two additional steps: click on the term selected, which automatically enters it into the search, and then press the "Search" button, as for any other matching search. Browses are performed when the user clicks on the appropriate button at the side of the drop-down menu. However, they are *not* recorded on the log; only matching searches are.

Each matching-search query is recorded on the log in full command language, even if the user does not actively input all of this language (it may be defaulted in the case of the basic or advanced search interface). Each search term is thus prefixed by a particular index label, multiple terms are separated by a Boolean operator (and/or/not), and also by any term-adjacency/proximity symbol; truncation symbols are recorded, as are any limiters entered as part of the query. A guide from the OCLC documentation showing the various, possible syntactic elements, together with examples, is shown below.

FIGURE 5.2 GUIDE TO FIRSTSEARCH COMMAND LANGUAGE

To search for...	*Use*	*Example*	*Results*
categories of information	index labels	su:sleepti:once	subject sleep, once in titles
exact phrase	quotes " "	"tunnel vision"	tunnel vision
plurals	plus sign +	plant+	plant, plants, plants', plant's, plantes, plantes'
variants or part of a word	wildcard *, #, or ?	zebu*	zebu, zebulon, zebutte
all words	AND	cold AND zinc	cold and zinc
one or all words	OR	cold OR zinc	cold, zinc, and cold and zinc
one word but not another	NOT	cold NOT weather	cold but not cold weather
words near each other, given order	w	cold w2 common	cold followed within 2 words by common
words near each other, any order	n	cold n3 common	cold and common, within 3 words of each other

Use an index label... and	*...when your search includes*	*Example*
colon (:)	individual words or fragments of phrases	kw:airline security
	search operators (*w* or *n*)	kw:alcohol w2 fetal
	other special search characters, such as the plus sign (+) or asterisk (*)	ti:(ocean+ or sea+) and pollut*
equal sign (=)	exact phrases as in names and titles	ti=alice in wonderland

5.10.2 Interpreting Query Elements

For this analysis, the coder is interested in identifying the search goal from each query logged. Certain elements of the logged search expression may not be related to search goal, but may instead be purely strategic. This may well include certain syntactic elements such as truncation. On the other hand, some syntactic elements, such as the index label, may well represent an aspect of the search goal that would otherwise be hard to identify, for example, whether the user's goal is to find documents from a particular journal title or documents on a particular subject, when the search terms themselves could be interpreted as either a journal title or a subject. The coder therefore needs to be clear how he is going to interpret each element of a query line: irrelevant to the search goal, or relevant, and if so, in what way? With these questions in mind, we shall discuss the various elements that might make up a query, as recorded on the log, beginning with the *syntactic* elements, before returning to the search terms themselves.

5.10.2.1 Index Labels

In the log, all search terms are prefixed by an index label. The same index label for the same kind of metadata element(s) is employed across databases. Labels whose indexes consist specifically of subject terms are assumed to prefix subject terms, except where the term is clearly meant to be a non-subject term (some subject vocabularies include non-subject terms). The following labels on FirstSearch are considered to represent subject indexes.

TABLE 5.7 LABELS TO REPRESENT SUBJECT INDEXES

Label	*Meaning*
ab:	Abstract
ag:	National Agricultural Library call number
bt:	Broad topics
bt=	Broad topics phrases
ca:	National Library of Canada call no.
cc:	Central concept
cd:	Classification descriptor

cd=	Classification descriptor phrase
ch:	Chemical indexing
ch=	Chemical indexing phrase
ci:	Citation
ci=	Citation phrase
cl:	Library call number
co:	Company name
co=	Company name phrase
cr:	Citation
cr=	Cited reference phrase
cs:	Chemical substance
cw=	Cited work phrase
dc:	Descriptor code
dd:	Dewey decimal call no.
de=	Descriptor phrase
gc=	Geographic coverage phrase
gn=	Geographic name phrase
id:	Identifier
id=	Identifier phrase
ii:	Identifier
lc:	Library of Congress call number
lm:	National Library of Medicine call no.
mc:	Major concept
mc=	Major concept phrase
mh:	MeSH heading
mh=	MeSH heading phrase
mj:	Major descriptors
mj=	Major descriptors phrase
mn:	Minor descriptors
mn=	Minor descriptors phrase
mt=	Marketing term phrase
na:	Named person
na=	Named person phrase
nc=	Named corporation phrase
nm:	Named person
nm=	Named person phrase
pc:	Product code
pc=	Product classification
pr=	Product name
sa=	Subject All
sh=	SubHeading (MeSH) phrase
si:	Industrial codes
sm=	Summary note
su:	Subject
su=	Subject phrase
tc=	Treatment Codes Phrase
ud:	Universal decimal class no.
us=	Geographic region
zq:	Subject area phrase

Labels whose indexes consist specifically of non-subject terms are assumed to prefix non-subject terms. All other index labels that were found on the logs were considered to fall into this category, with two important exceptions: the label for the general keyword index, which consists of both subject and non-subject terms, and the title keyword index, which might be used for subject searching, as well as known-item searching.

The general keyword label ("kw:") is also the most common index label on the logs. The keyword index includes subject descriptors that are wholly subject terms; it also includes title words and the like that *may* indicate subject, although these may also be used bibliographically. It does *not* include author names, however. Since most of the terms in the index would be either subject terms or quasi-subject terms, the coding will assume that keyword terms (defined by *kw:*) are subject terms, *unless* they lack any subject content.

Title index labels ("ti") also occurred quite frequently in the logs. Exact-phrase searches *(ti=)* or partial-phrase searches (using quotation marks) are presumed to represent genuine title searches. Even for basic title keyword searches (those with the *ti:* label and no quotation marks), the query is assumed to be of a known-work type, unless all the words (adjacent character combinations) prefixed by the title index label are grammatically disjointed such that the coder does not believe that the query is based on a specific title.

While in the vast majority of cases, terms are interpreted according to their labels, there were a small number of cases where terms clearly conflicted with their index labels. In such cases, provision was made for terms to be interpreted by the coder according to the meaning likely to be intended by the user. For example, a *content* word under the standard number index label is to be considered a subject word, as long as the context did not suggest otherwise.

5.10.2.2 Truncation

Some, though few, terms are operationally truncated on the log: both mid- and post-truncation is offered by FirstSearch. Truncation is interpreted here as a strategic measure only.

5.10.2.3 Boolean Operation

Boolean operators do not affect the decision as to whether a search is subject or non-subject or both. When they are changed in a modified search, this may be regarded as very likely caused by strategic considerations, and so such changes will not influence the coding.

5.10.2.4 Proximity Specification

Proximity symbols are very rarely recorded in the logs. The symbol "w" or "w1" suggests a phrase; all words involving *w* symbols are to be treated as phrases where a phrase can be made out by the coder. Words involving "n" symbols, on the other hand (which mean "near" words but in any order), are to be treated as grammatically disjointed words.

5.10.2.5 Search Limiters

Those fields presented as search "limiting" options in FirstSearch are in fact recorded in the same way as are other fields, by the index label. For example, when the full-text limiter is "on," this is recorded in the log as: "and ft: fulltext". Search limiting – by format, language, etc. – might relate to a goal or revised goal, and so in most cases it will be coded a such; however, we should bear in mind that in many cases it is probably used strategically, and in the refined index, codes pertaining to search limitation will be discounted. Further, there are two limiting fields that will be ignored by the coder from the outset, namely, those for full-text availability ("ft:") and library availability ("li:"). Both of these limiters are practical ones and cannot be interpreted as part of the substantive search goal – all search goals require the availability of relevant documents (unless the

searcher wishes to confirm their unavailability); whether the searcher elects to limit to those presumably more readily available is related to the searcher's external environment and their search situation, generally not to the search goal itself.

5.10.2.6 Search Terms

Each word or other combination of adjacent characters (perhaps a number or code), which makes up a query line would have been treated separately by the FirstSearch system unless quotation marks or proximity symbols are incorporated. The system would combine such words with the default Boolean operator, AND.

What the system does, as was noted above, is not necessarily what the user might have intended. It may well be, for example, that the user intends a phrase consisting of more than one word, but sees no need to expend additional time and energy informing the system of this – results may be good enough without such expenditure. The question is, how should the coder interpret the user's intention when encountering adjacent words in a query that possess meaning on their own, but can also be grouped into phrases. We can only make a general rule for the coder to follow, such that he is consistent across databases. Such rules are based on what is perceived to be the most likely intention.

As stated above (section 5.10.2.1), search words under the title keyword index label are to be examined together for any indication of a phrase. On the other hand, the rule for combinations of words under the keyword index label and those words deemed as pertaining to subject (normally indicated as such by an index label included in the list in section 5.10.2.1 above), is that if phrases present themselves as likely representations of user intention, then they are interpreted as a single search term, but that if it is not clear whether a phrase or individual words are intended, given the context of the session, then the individual words are each treated as search terms. All non-subject elements are treated separately, bounded by their index label.

5.10.3 Subject versus Known-Item Queries

Only those queries pertaining to subject searches are to be analysed in this study. The next step, then, is to provide the coder with guidance on how to determine which queries do not pertain to subject searching and thus can be coded as such.

It was decided that a subject search might include non-subject aspects. Non-subject elements may be included in a query for practical, strategic or objective reasons. A practical reason may be, for instance, that the searcher wishes to limit by a particular library, since that is the one he is in. The non-subject element may be strategic – for example, the user limits by date because they fear too much recall otherwise. Or the element may be included because the searcher's goal really is not just for documents on a particular subject, but for those in a particular format, language, etc.

However, some kinds of non-subject element, such as an ISBN, smack so much of a known-item search that a query containing one will be counted as a non-subject query regardless of any other "subject-looking" elements that it might include. For the purposes of this analysis, we shall not use the term "known-item search" in a very literal sense – it shall embrace known-work searches, known-author searches, known-series searches, indeed searches on any specific, non-subject entity, or group of non-subject entities.

What we need to predetermine for the coder, therefore, are those query elements that usually indicate a known-item search, so that their queries can be coded as ones not pertaining to subject. The following index labels on the log are considered to be for non-subject indexes which indicate the likelihood of a known-item search.

TABLE 5.8 LABELS TO INDICATE NON-SUBJECT SEARCH

Label	*Meaning*
aa:	Author affiliation
au:	Author
au=	Author phrase
cf=	Named conference phrase
cn:	Conference name
cn=	Conference name phrase
ea:	Extended author(s)
ea=	Extended author(s) phrase
ed:	Edition date
ed=	Editor phrase
ib:	ISBN
in:	Institution
in=	Institution phrase
jn:	Journal name
jo:	Journal title
jo=	Journal title phrase
nb:	ISBN
oc:	OCLC number
pb:	Publisher
pb=	Publisher name phrase
pb=	Publisher phrase
pd:	Publication date
pk:	Publisher name
pk=	Publisher name phrase
pl:	Place of publication
pl=	Publisher location
pn:	Personal name
pn=	Personal name phrase
rf:	Reference
se:	Series
se=	Series phrase
sn:	Standard number
so:	Source
so=	Source phrase

Any query which includes one or more of the above types of search term, will be coded as a known-item query. Further, any query which does not include what the coder interprets as a subject term, will also be coded as a known-item query.

On the other hand, when other non-subject elements, such as those for format and language, are combined in a query that contains at least one subject term (according to the coder's interpretation), that query will be deemed a subject query, and coded as such.

5.10.4 Coding of Subject Queries

The query lines in each logged session are in chronological order. The coding will thus proceed through the lines of a session chronologically. The first query identified as a subject search (as defined above), is regarded as an *initial subject query.* The subsequent coding is intended to identify redefinition. Thus queries within the same session that appear to be a *following on* from the initial query should be coded in relation to this initial query, as a type of *modified query.* On the other hand, subsequent subject searches that appear to bear no semantic resemblance to the initial query, should instead be coded independent of it. We have established in section 5.3.4 that this non-resemblance will be defined as such when *no* topical term is semantically related or identical to a topical term contained in the initial query. Conceptually unrelated queries are assumed to be based on separate search goals and not a goal that is the same or a modification of the one on which the initial query is based. It is, of course, entirely possible that a user should search for items on two or more different subjects within the same session.

A "new" subject query, unrelated to the initial query, may also lead to redefinition, the subject of this study, and is to be coded as another initial subject query. In this way, a session may consist of multiple initial queries and multiple "families" of queries based on their respective initial queries.

There is one slight complication that might occasionally arise with respect to the identification of new versus modified queries. There may be cases when a user embarks on a new search, with a different search goal, but then later reverts to their prior search. In other words, a query may be related to an initial query that is not the last initial query chronologically (within the same session). Thus, the coder must examine the relationship (if any) between a given query and all prior initial searches within a session. If a query could be related to more than one initial query (this would be unlikely due to their lack of relationship, but a query might conceivably fall "in between"), then the policy adopted for the coding is to assume that it is related to the latest initial query.

Not all subject queries will be modifications or new initial queries. Some may be exact repetitions of previous queries, others may differ from previous queries only with respect to a synonym or near-synonym and/or syntax. In both cases, it is considered most likely that the query was entered for strategic reasons and thus for the purposes of this analysis not a modification that might be the result of a revised search goal. As such, this kind of query, to be called a "repeat query," will be assigned a separate code of its own, outside of those assigned for modified and initial queries. How the coder might judge as to a new term's synonymity with an old term is discussed in section 5.3.4. If in doubt, the coder should treat a new term as a semantic modification, rather than a synonym. We may add here that the coder is to treat changes of spelling and form (e.g. singular to plural) of a term as equivalent to synonymity.

5.10.5 Categories of Modified Queries

As was indicated in section 5.3.5 above, a matrix containing nine distinct categories of modified query was established for this analysis. This represents the different combinations of two aspects of query modification, in relation to an initial query: whether modification involves topical or non-topical elements, or both; and whether the modification involves the introduction of new conceptual elements, or the elimination of original elements, or both.

Although the second aspect is represented in this coding system by the concept of "broadening" and "narrowing" of searches, the definitions of "broader" and "narrower" queries are *not* intended to cover all kinds of broadening and narrowing in the information search context. For example, subject terms may be replaced by related terms which are higher or lower in a topical hierarchy. This kind of broadening or narrowing may or may not be related to a revised search goal, but in any case, it is not covered in the coding system developed for this analysis. The reasons for this are that a term's semantically hierarchical position, in relation to another term, would sometimes be very hard for the coder to call, and that even if it were categorized, it is by no means clear that such a categorization would be useful when attempting to distinguish between redefinition-induced query modifications and strategy-induced ones.

5.11 Code Definitions

Guidance was given to the coder as detailed above. Reproduced below are the definitions of the twelve codes that could be assigned to each query, which were used by the coder throughout the analysis; these definitions had been developed following experimentation and parallel coding (to check reliability). Only one code may be assigned a query, so the definitions are written in such a way as to make the codes mutually exclusive.

K = known-item query

Query *includes* standard number, author, series, or publisher/place index label, or is a title browse, or consists only of "non-subject" terms (i.e. that lack subject content). For queries including title keywords, judgement is exercised as to whether the user has entered a known title or part of a title, or is conducting a subject search via the title index. If some or all of the title words represent a phrase, or a phrase minus non-content words such as prepositions, then a known-item query is presumed.

A name or a single word may or may not be coded as a known-item query depending on the context of the rest of the search session – if in doubt such queries are regarded as subject queries.

K searches take precedence over all of the following query codes.

S = initial subject query

Query is not classed as *K* above, and includes subject, title or general keyword index label; it may be qualified in some way such as by material or language or date or location/source, and such qualifiers may include non-subject terms entered as general keywords, etc. Another query with a different subject goal also logged within a session, will likewise be coded as an *S* query, defined as a query which bears no subject-content relation to previous *S* queries in the same session, i.e. none of the earlier subject terms are represented in the search, nor does

it include any near-synonym, nor indeed could any of its terms be construed as being semantically related.

X = repeat query

Defined in relation to a previous *S* or consequential search in the same session, it is exactly the same query, or is different only with respect to a modification of index labels, and/or the operation of truncation, proximity, and/or Boolean logic, and/or the presence or absence of the library location and/or full-text limiter, and/or replaces one or more search terms with exact (or almost exact) synonyms (such that the user is unlikely to have differentiated semantically), including those with different spelling or form.

Modified queries

The following query codes relate to a previous *S* query in the same session (and normally to the last *S* query). In cases where more than one *S* query may apply, it is assumed to relate to the latest one. In their definitions, "non-subject terms" exclude those for library location and full-text limiter, which are to be ignored. Modifications as listed under query type *X* above, are also ignored. The table below gives the code for each combination of difference in subject and non-subject terms.

TABLE 5.9 CODES FOR COMBINATIONS OF DIFFERENCES IN SUBJECT AND NON-SUBJECT TERMS

	f-	*f*=	*f*+	*fx*
su -	BC	BS	MC	MT
su =	BT	X	NT	MT
su +	MC	NS	NC	MC
su x	MS	MS	MC	MC
su %	MS	MS	MC	MC

Key

f non-subject term (except library location or full-text limiter) identified as such
su subject term identified as such
su - elimination of one or more original subject terms which have not been substituted with a semantically related term
su = no addition of new terms or elimination of original subject terms
su + addition of one or more subject terms which are not semantically related substitutes for original terms
su x elimination of one or more original subject terms which have not been substituted with a semantically related term *and* addition of one or more subject terms which are not semantically related substitutes for original terms
su % substitution of an original subject term with a semantically related but not synonymous term, irrespective of elimination and/or addition of other subject terms
f- elimination of one or more original non-subject terms
f= no addition of new non-subject terms or elimination of original non-subject terms
f+ addition of one or more non-subject terms
fx elimination of one or more non-subject terms *and* addition of one or more new non-subject terms

NS = narrowing of initial query in terms of subject

Query contains the same terms and logic except for one or more additional subject terms which are not semantically related substitutes for original terms.

NT = narrowing of initial query in non-subject terms

Query contains the same terms and logic except for one or more additional non-subject terms.

NC = narrowing of initial query in both subject and non-subject terms

Query contains the same terms and logic except for one or more additional subject terms and one or more non-subject terms, which are not semantically related substitutes for original terms.

BS = broadening of initial query in subject terms

Query contains the same terms and logic except for one or more, fewer subject terms which have not been substituted with a semantically related term.

BT = broadening of initial query in non-subject terms

Query contains the same terms and logic except for one or more, fewer non-subject terms.

BC = broadening of initial query in both subject and non-subject terms

Query contains the same terms and logic except for one or more, fewer subject terms and one or more, fewer non-subject terms, none of which has been substituted with a semantically related term.

MS = mixed modification of subject terms

Query with at least one semantically related but not synonymous subject term replacing an original term, but with no new non-subject terms.

MT = mixed modification of non-subject terms

Query with at least one new non-subject term and in which at least one original non-subject term has been eliminated, but with no new subject terms.

MC = mixed modification of both subject and non-subject terms

Query with at least one semantically related but not synonymous subject term replacing an original term, and with at least one new non-subject term, and in which at least one original non-subject term has been eliminated; or where one or more subject terms have taken the place of one or more non-subject terms, or vice-versa.

5.12 Parallel Coding (Inter-Coder Agreement)

Given the subjective component involved in the coding of the logs, it was decided to perform preliminary parallel coding on two specimen log sections, one comprising WorldCat-only sessions, the other, sessions involving one or more full-text downloadings. Both of the specimen logs consisted of 250 lines. The author and a colleague thus coded the same specimen logs; the author was to carry out all of the coding for the analysis. The colleague also had a librarianship background and was asked to approach the coding from a "general knowledge" background, regardless of topics indicated on the log.

The author did not discuss specific examples with the parallel coder beforehand, but did run through the written definitions of codes, which were the prototypical version of those reproduced in section 5.11 above. The two specimen log sections were coded over the same week, in the same order, without discussion. The resulting codes were then compared.

The percentage of queries coded the same, out of the dozen codes available (at least after an initial search), was 91% and 96% for the two sets. The Kappa coefficient, sometimes used (though controversially) for inter-rater agreement, for the full-text and WorldCat sets was 0.75 and 0.79 respectively, which would indicate "good" agreement. The cases of disagreement were examined and found to be mostly *errors,* as opposed to differences of interpretation – in such cases, the coder "in error" readily admitted this in the post-mortem. However, a few differences of interpretation were found and as a result the definitions for some of the codes were clarified and/or expanded (particularly those elements concerning title index searches). One or two of the errors were *compound* errors, resulting from a divergence in a previous search.

Given that the parallel coding had revealed no major areas of disagreement that might be due to subjectivity and that the levels of agreement looked as though they would produce similar findings throughout this analysis, the author's coding was considered reliable enough to proceed. An extract of the coding is provided in Appendix B.

5.13 Two Indices of (Strong) Redefinition

During the coding of the logs, the nature of the differently coded query modifications was examined, as were the patterns that they involved. Additional non-subject terms were seen as more often being used for strategic purposes than were subject terms. It may be speculated that this might be because specific subject terms are more difficult to think of, whereas the searcher can often recycle stock non-subject terms for strategic purposes. Where subject terms have simply been deleted from a query in order to broaden a search, it appeared to the author that this broadening was often intended to produce higher recall than because of a genuine reconceptualisation of the search goal as a broader topic. More commonly, it is supposed, will a search goal become more specific, rather than less so.

Following these observations, it was hypothesised that the two query codes most likely to represent goal redefinition, as opposed to strategy, were *NS* and *MS* (see section 5.11). An index based on all nine codes of query modification was thus supplemented by another index based on just the *NS* and *MS* codes, taking account of only those modifications with substituted or additional subject search terms.

These two indices of "strong" search redefinition (as defined in chapter 4), were defined as the number of modified queries stemming from each initial subject query. The index embracing all nine modification codes was to be known as the "raw" index, while that using the two codes, as the "refined" index.

It is assumed that both indices would reflect search redefinition *uniformly* across databases, but it is hypothesised that the refined index would be a "purer" measure of redefinition and as such would show similar results, but ones more pronounced, in the sense that the degree of significance or non-significance would be higher than in the case of the raw index, which would be "blunted" by more "noise" caused by strategic query modification.

Frequency distributions based on both indices were produced for each of the thirteen datasets listed in section 5.8.4 above, as well as for two control datasets.

5.14 Examination of a Metadata Content Factor

5.14.1 Abstract Frequency and Length

Once the datasets for analysis had been finalised, the author reviewed the statistical analyses that had been planned, and concluded that detailed comparisons of datasets in relation to the (displayed) metadata elements covered section 5.8.1 above, were hampered by a lack of representation in some of the metadata groups, particularly groups A-C. What did show promise, however, was an analysis of a more general way in which the potentially largest "chunk" of metadata might affect redefinition, namely, abstracts. Thus a more detailed study of the frequency and length of abstracts in the relevant FirstSearch databases was undertaken.

The initial survey of abstract presence (see section 5.8.1 above) had been based on "typical" records. A deeper survey was necessary in order to refine, if necessary, the trichotomous variable of yes/some/no. To do this, "typical" queries (typical in relation to those being analysed in the logs) were simulated in order to discover the frequency with which abstracts would be available to the user. It was realised at this point that abstracts were likely to be more common than initially thought, since general keyword queries (the typical type of query) would retrieve records with abstracts much more readily than they would records without abstracts – abstract words, when present in a record, generally accounted for a large proportion of a record's entries on the keyword index.

A sample of subject queries across databases showed a median hit rate of around 10 records, which diluted the problem of the default ordering of results, which, being reserve accession date, approximated to descending publication date. If the median hit rate had amounted to much more than a screen's worth of results, there would likely be more brief records at the tail end of result sets not seen by the user, and the non-presentation of such records would have an affect on

average abstract frequency in a way that might vary across databases given that databases had different histories of abstracting and different practices of updating records with abstracts.

The queries to be used for generating the sample result sets employed were derived from a random selection of those found in the logs, where result sets of 100-120 records were produced – samples of over 100 records would have a standard error of less than 5%. The number of records with abstracts, in these result sets was counted. Estimates of the typical frequency of abstracts for each dataset were based on these counts, and are shown below.

Dataset	Abstract frequency (% of records)
Group B	0
Group Z	0
OCLC ArticleFirst	0
ABI/INFORM	100
Business Management	90
Education Abstracts	90
Health & Wellness	50
Medline	95
Periodical Abstracts	100
Sociological Abstracts	95

TABLE 5.10 ABSTRACT FREQUENCY

The above findings indicated that typical queries in most of the datasets would likely yield either very high levels of abstracts or very low levels. The one exception was the Health & Wellness dataset, where a mid-level was suggested.

A distinction between 90% and 100% frequencies found in the above table was considered unreliable, especially given the fact that the actual frequency depended not only on the latent frequency of abstracts, but which full records users displayed. Thus the trichotomous scale previously established was retained: high (including complete), medium, and low (i.e. none).

Dataset	Abstract frequency
Group B	Low
Group Z	Low
OCLC ArticleFirst	Low
ABI/INFORM	High
Business Management	High
Education Abstracts	High
Health & Wellness	Medium
Medline	High
Periodical Abstracts	High
Sociological Abstracts	High

TABLE 5.11 ABSTRACT FREQUENCY (COLLAPSED)

It was observed that some abstracts in FirstSearch are considerably longer than others. The question then arises: could an abstract of, say, fifty words be significantly less likely to yield redefinition than one of two hundred words? Or is there a law of diminishing returns applicable here – after the searcher reads a certain number of sentences, if redefinition were to happen, then it most likely would have happened by then? The answer is unclear, and so worth testing. It was determined that those databases with a high level of abstracts would be further examined and a mean length of abstract estimated, using the result sets derived for the frequency count above. The estimates are shown in table 5.12 below.

Database	Mean abstract length (words)
ABI/INFORM	75
Business Management	50
Education Abstracts	80
Medline	130
Periodical Abstracts	30
Sociological Abstracts	120

TABLE 5.12 ABSTRACT LENGTH (AVERAGE)

Again, it was thought that some of the distances between these figures were not large enough to be reliable for ranking purposes, particularly since the particular

abstracts encountered by the searcher would vary in length considerably; it could also be argued that the median would be a better measure of "typical" length encountered than would the mean. Thus the estimations were collapsed into a broader, trichotomous classification as below.

Database	Abstract length
Business Management	Low
Periodical Abstracts	Low
ABI/INFORM	Medium
Education Abstracts	Medium
Medline	High
Sociological Abstracts	High

TABLE 5.13 ABSTRACT LENGTH (COLLAPSED)

Bilateral comparisons could have been made between the redefinition indices produced by the various datasets in tables 5.11 and 5.13 above, but the assumption of control over other possible factors for such comparisons was unconvincing, and instead a multilateral test *combining* the specific datasets in tables 5.11 and 5.13, was undertaken. The test was based on the *relative* amounts of redistribution (according to the raw and refined indices) produced by the datasets and their relative abstract frequencies and lengths, that is, on rankings of the datasets for redefinition and for abstract frequency and length. Clearly the datasets needed to be *ranked,* since any correlation between a redefinition index and abstract frequency or length would very likely be indirect – the association between the variables would be nonparametric. Tables 5.11 and 5.13 were used for abstract frequency and length rankings. For the redefinition ranking, the *central tendencies* of the raw and refined indices for each dataset needed to be established.

The median was selected as the primary measure of central tendency, given the skewness of the distributions of the redefinition indices and the danger of outliers having been produced by atypical searching behaviour independent of an abstract factor. Since this resulted in many ties in the ranking, the inter-quartile mean was used as a secondary measure. Since there were still ties in the rankings of abstract frequency and length, Kendall's *tau b* correlation coefficient (as opposed to

Spearman's *rho*) was chosen for the significance tests, the null hypothesis being that there is no correlation between variables.

Comparison between redefinition indices for the Group Z dataset and the dataset representing sessions with full-text downloading as a whole (dataset 3), could also show whether the *absence* of abstracts was a factor on redefinition. This assumes that the Group Z dataset does not represent other possible redefinition factors significantly more or less so than does dataset 3, an assumption which was considered reasonable in this case.

This comparison could be carried out on the distributions of the indices, or on the distributions' central tendencies. It would have been more intuitive to compare central tendencies, so that the test would have directly supported or rejected a hypothesis that one dataset represented more redefinition, on the *average,* than the other. The distributions of query modification were *not* assumed to be parametric – they may approximate to a particular distribution (logarithmic, for example), but there appear to be no theoretical grounds for assuming so. Thus a non-parametric test was sought. However, the nonparametric Mann-Whitney test for location (that is, on the medians) was not used, since (a) the author could not be sure if the distributions would approximate to the same shape, and (b) it is not reliable in the case of many tied observations, as was the case here. The author therefore resorted to comparison between distributions, using the Kolmogorov-Smirnov 2-sample test. This non-parametric test indicates whether a given two samples are unlikely to be derived from the same population, in other words, if there is a significant difference between the sample distributions.

5.14.2 Presence of Descriptors and/or Identifiers

Although a detailed survey of the possible effect of descriptor and identifiers was not considered feasible given the range of datasets available to the author for comparison, it was possible to gain an indication of the relative impact of descriptors and indicators on search redefinition by comparing the redefinition indices of the Group Z dataset, which represented searching on databases without abstracts, but with descriptors and/or indicators, with those of the OCLC

ArticleFirst dataset, which represented searching on a database with no abstracts, but also no descriptors or identifiers. The same Kolmogorov-Smirnov test was used for this comparison.

The author also thought it worthwhile to compare the indices for OCLC ArticleFirst, which includes no abstracts, descriptors or identifiers, with those of the dataset representing searches on other full-text databases (dataset 3). This would confirm, or reject, the effect of the absence of the three metadata elements as a whole.

5.15 Examination of Other Factors

5.15.1 Bibliographic versus Full-Text Searching and Item versus Article Searching

Datasets 1-3 provided the author the opportunity to make two other useful comparisons between dataset results, again employing the Kolmogorov-Smirnov test to indicate significant difference or otherwise. The test was used to compare redefinition on WorldCat versus other non-full-text databases (or at least databases which did not induce full-text downloadings); and on full-text versus non-full-text databases (or at least databases which did not induce full-text downloadings).

5.15.2 Hit Rate Factor

Whether a low hit rate is likely to impact on the extent of redefinition is to be investigated in two ways. First, the proportion of initial queries producing zero hits, less than six and less than eleven hits was calculated for each dataset, 1-13. These proportions were ranked and compared with the raw and refined index rankings for the same datasets. Kendall's *tau b* correlation coefficient was calculated. Second, a more sophisticated analysis was performed, whereby the number of hits for initial queries in a section of the Medline dataset was compared with their respective number of modified queries (in both raw and refined index terms). Only those initial queries with ten or fewer hits were

included in this analysis. Kendall's *tau b* correlation coefficient was again applied.

5.15.3 Retrievability Factor

As discussed above in section 5.7.4, we need to test whether *retrievability* of abstracts had any significant effect on redefinition. We hypothesized earlier (in section 5.7.4) that greater retrievability of records through more index entries leads to a lack of precision which, in turn, leads to more search redefinition.

In order to test this "retrievability" factor, comparison was made between *kw* and *su* searches recorded in the WorldCat dataset, which was chosen because few *su* searches were available for analysis from any of the other datasets. The proportion of queries following a purely subject *kw* query which were coded as modified queries (according to raw and refined indices) was compared with the proportion of queries following a purely *su* query which were coded as modified queries. Fisher's Exact test was used to test the null hypothesis that there is association between the two proportions, that is, additional redefinition results from greater retrievability of records. (We are assuming here that the result with respect to immediately subsequent queries also applies to other subsequent queries.)

5.15.4 Subject/Discipline Factor

Many of the datasets featured in the cross-database analyses represented a broad range of subjects and disciplines, and were in this respect controlled. Six datasets, however, which featured in the correlation tests for the abstract factor, represented databases which were geared towards particular disciplinary audiences. The discipline area of each of these databases was determined in accordance with the OCLC documentation, and is given below.

	discipline area (DDC)
ABI/INFORM	330
Business Management	330
Education Abstracts	370
Health & Wellness	610
Medline	610
Sociological Abstracts	300

TABLE 5.14 DATABASE DISCIPLINE AREAS

Given the chance of subject/discipline affecting the correlation tests, it was necessary to test for association between discipline and redefinition. This could be done given the clustering of the six databases: three groups of two could be identified, with Education Abstracts and Sociological Abstracts paired on the basis of their discipline areas' relatively similar nature. Using the refined index to rank the datasets according to redefinition, a Kruskal-Wallis test was performed to test for any association between discipline and redefinition; given the low number of cases, direct observation could also be used here. A result which indicated no particular association would not rule out some effect that some disciplines might have on redefinition, but what it would do is substantiate the validity of the results of the correlation tests for the (possible) abstract factor.

5.16 Additional Control Tests

The Kolmogorov-Smirnov 2-sample test was also employed to check the validity of the other Kolmogorov-Smirnov tests based on the two indices of redefinition. Distributions of the indices derived from two sections of the WorldCat dataset were compared using the K-S test. The two sections represented two different days of search sessions. The analysis was then replicated using the Medline dataset.

Given the apparent lack of possible factors that might affect differences of redefinition in these comparisons, it was supposed that the K-S tests would *not* reveal any particular differences between the pairs of distributions. If this were the case, while the other K-S tests *did* reveal differences, then the redefinition

indices would likely show *some* difference in searching behaviour due to interaction with different FirstSearch databases, although this difference might not necessarily be pertaining to search goal redefinition alone.

Chapter 6 RESULTS

6.1 Indices of Redefinition

6.1.1 Distributions

As explained in chapter 5, two indices of search redefinition were constructed according to the number of modified queries stemming from each initial subject query coded from the logs. For both raw and refined indices, the number of modified queries derived from each initial query was tabulated to produce frequency distributions for each of the thirteen datasets listed in section 5.8.4 above. These distributions are detailed in Appendix C (i) and C (ii).

In addition, distributions of the raw and refined indices were constructed for two pairs of control datasets, derived from the WorldCat and Medline datasets. These distributions are detailed in Appendix C (iii).

6.1.2 Central Tendencies

The central tendencies of the distributions of the redefinition indices for the thirteen main datasets are shown in table 6.1 below, in terms of their medians and inter-quartile means.

Dataset	Sample size		Raw index		Refined index	
	raw	*refined*	*median*	*iqmean*	*median*	*iqmean*
WorldCat	433	433	0	0.36	0	0.03
Non-full-text	627	627	0	0.41	0	0.27
Full-text	287	287	1	1.03	1	0.81
Group B	12	12	1	1.83	0.5	1
Group Z	65	65	1	1.12	0	0.7
OCLC ArticleFirst	22	20	0	0.5	0	0.4
ABI/INFORM	26	26	1	1.29	1	1.07
Business & Mgt	54	45	2	1.86	2	1.65
Education Abstracts	36	28	1	1.11	0.5	0.71
Health & Wellness	216	190	1	1.37	1	1
Medline	161	142	1	1.09	1	0.74
Periodical Abstracts	40	38	1	1.2	0.5	0.75
Sociological Abstracts	32	27	1	1.5	1	1.27

TABLE 6.1 CENTRAL TENDENCIES FOR DISTRIBUTIONS

We may observe a reasonable correlation between the two measures of central tendency. The median was considered the theoretically most suitable measure, but given the ties, the interquartile means were calculated as a secondary measure.

It is interesting to compare these distributions with that found by Jansen, Spink and Saracevic (2000:212), who identified query modifications of any kind following initial searches (which they termed unique searches). The mean of 1.6 modifications per unique search and a Zipf type of distribution indicates that while Web search engines may be used more briefly than some of the more "old-fashioned" OPAC and IR systems, query modification through Excite does not appear to be all that different from that through FirstSearch, at least not in terms of quantity.

6.1.3 Control Tests

Prior to analysis of the differences between the redefinition indices produced by different datasets, the distributions for the control samples were examined, as per section 5.16 above. The results of the Kolmorogov-Smirnov test for distributional difference are shown below in table 6.2.

Database	Raw index		Refined index	
	D-statistic	*p-value*	*D-statistic*	*p-value*
WorldCat	0.0811	0.781	0.1105	0.398
Medline	0.1566	0.876	0.1756	0.520

TABLE 6.2 K-S TEST FOR CONTROL DISTRIBUTIONS

With mid-high *p* values, the null hypothesis of no difference between the pairs of distributions was not rejected. We should note that despite high *p* values, the sizes of the control samples were within the range of the other sample sizes, and so the power would not have been particularly low. Assuming that search behaviour on different days but on the same database in likely to be similar, this result indicates that testing on these redefinition indices is unlikely to produce a

false positive, that is, it does not show significant differences between datasets that are not actually present.

If we assume that the indices from two datasets would be different due to differences in the amount of search redefinition and not because of differences in search strategy, then these control tests provide evidence that should a statistically significant difference between the indices from two datasets be detected, then this difference would represent a difference in the amount of search redefinition. The reason for such a difference is another matter.

6.2 Metadata Effect

We have hypothesized that differences in metadata content affect the extent of search redefinition. This may particularly the case with respect to important types of metadata such as abstracts, descriptors and identifiers.

The first test to find out whether this is in fact the case involved comparing the index distributions for OCLC ArticleFirst, which includes no abstracts, descriptors or identifiers, and dataset 3, which represents searches on the other full-text databases, many of which would include one or more of these metadata elements. The results of the test are shown below: a significant difference was detected.

Sample pair	**Raw index**		**Refined index**	
	D-statistic	*p-value*	*D-statistic*	*p-value*
Full-text				
OCLC ArticleFirst	0.5876	<0.001	0.6047	<0.001

TABLE 6.3 K-S TEST FOR METADATA FACTOR

6.2.1 Abstract Frequency

To analyse whether or not abstract frequency affects search redefinition, another Kolmogorov-Smirnov test was carried out on the difference between index distributions for the Group Z dataset, which represents searching on databases

with descriptors and identifiers, but not abstracts, and those for dataset 3, which represents searches on the other full-text databases, many of which would include abstracts. The results below show a significant difference, which suggests that the *presence* of abstracts, at least, has an effect on search redefinition.

Sample pair	Raw index		Refined index	
	D-statistic	*p-value*	*D-statistic*	*p-value*
Full-text				
Group Z	0.3213	<0.001	0.4046	<0.001

TABLE 6.4 K-S TEST FOR ABSTRACT FACTOR

Evidence of a greater effect was sought by performing the correlation test detailed in section 5.14.1. Both the raw and refined indices of the datasets featured in table 6.5 below were ranked according to their respective medians in the first instance, and according to their interquartile means in the second. The datasets were also ranked according to abstract frequency, as per table 5.11.

Database	Rank		
	raw index	*refined index*	*abstracts*
Group B	2	6	9
Group Z	7	9	9
OCLC ArticleFirst	10	10	9
ABI/INFORM	5	3	3.5
Business Management	1	1	3.5
Education Abstracts	8	8	3.5
Health & Wellness	4	4	7
Medline	9	5	3.5
Periodical Abstracts	6	7	3.5
Sociological Abstracts	3	2	3.5

TABLE 6.5 RANKING OF TEN DATABASES

The degree of correlation between the index rankings and the abstract frequency rankings was tested using Kendall's *tau b* correlation coefficient, the results shown in table 6.6 below.

	abstracts
distribution (raw)	0.163 (p=0.569)
distribution (refined)	0.373 (p=0.178)

TABLE 6.6 COEFFICIENTS AND *P* VALUES FOR ABSTRACT FREQUENCY

The *p* values, shown in parentheses above, are two-tailed. They show no particular evidence of a relationship between abstract frequency and the indices of redefinition. However, the results do not necessarily contradict the results of the K-S test shown in table 6.4 which pointed to redefinition induced by the presence of abstracts. The refined index was positively correlated with abstract frequency, with a *p* value of 0.178, suggesting that abstracts might still be a factor on search redefinition, though perhaps not a strong one. In fact, this *p* value is reduced to a value verging on the 0.1 level of significance if a different measure of central tendency and/or a dichotomous frequency scale is used.

6.2.2 Abstract length

The six datasets with a high level of abstract frequency were ranked according to their indices' respective medians (in the first instance) and interquartile means (in the second), and according to their abstracts' length, as per table 5.13.

Database	**Rank**		
	raw index	*refined index*	*abstract length*
ABI/INFORM	3	3	3.5
Business Management	1	1	5.5
Education Abstracts	5	6	3.5
Medline	6	4	1.5
Periodical Abstracts	4	5	5.5
Sociological Abstracts	2	2	1.5

TABLE 6.7 RANKING OF SIX DATABASES

The degree of correlation between the index rankings and the abstract length rankings was tested using Kendall's *tau b* correlation coefficient; the results are shown below.

	abstract length
distribution (raw)	-0.298 (p=0.427)
distribution (refined)	0.333 (p=0.348)

TABLE 6.8 COEFFICIENTS AND *P* VALUES FOR ABSTRACT LENGTH

The *p* values, shown in parentheses above, are two-tailed. They show no particular evidence of a relationship between abstract length and the indices of redefinition. We should note that the correlation test might be less reliable where there are only six cases, in relation to abstract length, but direct observation supports the high *p* values. While abstract presence may be a factor on redefinition, the length of an abstract might be of little significance, for more words may give diminishing returns. The user may spend very little more time, on the average, reading a long abstract than reading a short one.

The evidence points to an abstract factor, but not a very strong one. It may well be that other factors are more significant.

6.2.3 Descriptors and Identifiers

Another type of metadata effect that might be significant is the presence or absence of descriptors and identifiers. The index distributions of the Group Z dataset with those of the OCLC ArticleFirst dataset were compared by means of the Kolmogorov-Smirnov test; the results are shown below.

Sample pair	Raw index		Refined index	
	D-statistic	*p-value*	*D-statistic*	*p-value*
Group Z				
OCLC ArticleFirst	0.421	0.004	0.45	0.003

TABLE 6.9 K-S TEST FOR DESCRIPTOR/IDENTIFIER FACTOR

The results show a significant difference, which suggests that the presence of descriptors and/or abstracts is a factor. It is speculated that it is the hyperlinked nature of the descriptors which might be the main cause of redefinition here. Further investigation is required.

6.3 Bibliographic versus Full-Text Searching

The difference between searching on databases which offered full text and searching on those that did not (at least not to the user) was tested by comparing the redefinition indices for dataset 2 (no full-text downloading) and dataset 3 (full-text downloading). Again, the Kolmogorov-Smirnov test was used.

Sample pair	Raw index		Refined index	
	D-statistic	*p-value*	*D-statistic*	*p-value*
Non-full-text				
Full-text	0.2324	<0.001	0.2494	<0.001

TABLE 6.10 K-S TEST FOR DOCUMENT TYPE FACTOR

The results show a significant difference in redefinition, with subject searching on full-text databases likely to result in more redefinition than on those where full text is not available.

6.4 Item versus Article Searching

The Kolmogorov-Smirnov test was also used to compare redefinition in the WorldCat dataset with that found in dataset 2, which represented searching on other databases without full-text downloading. The dataset 2 databases would

generally produce records for articles; the WorldCat database contained records for item-level documents, in the main.

Sample pair	Raw index		Refined index	
	D-statistic	*p-value*	*D-statistic*	*p-value*
WorldCat				
Non-full-text	0.1765	<0.001	0.2279	<0.001

TABLE 6.11 K-S TEST FOR FULL-TEXT FACTOR

The results show a significant difference in redefinition, with subject searches on WorldCat less likely to produce redefinition than subject searches on periodical indexes.

6.5 Hit Rate

The proportion of initial queries producing zero hits, less than six and less than eleven hits was calculated for the datasets 1-13. The proportions are shown in table 6.12 as percentages and ranks. The table also shows the rankings for redefinition based on the medians and interquartile means of the refined and raw index distributions.

Dataset	*0 hits*		*<6 hits*		*<11 hits*		*redefinition (raw)*	*redefinition (refined)*
	%	*rank*	%	*rank*	%	*rank*		
WorldCat	15.5	4	36.8	5	44	6	13	13
Non-full-text	29.6	12	48.5	12	54.1	12	12	12
Full-text	24	8	43.2	8	51.3	9	10	5
Descriptors only	25	9	37.5	6	37.5	3	2	7
Non-abstract	22.5	7	38.5	7	47.7	7	7	10
OCLC ArticleFirst	0	1.5	0	1	0	1	11	11
ABI/INFORM	8.7	3	30.4	3	39.1	4	5	3
Business & Mgt	15.6	5	34.4	4	40.6	5	1	1
Education Abstracts	0	1.5	0.1	2	0.2	2	8	9
Health & Wellness	20.6	6	44.7	9	50.9	8	4	4
Medline	29.2	11	45.3	10	52.1	10	9	6
Periodical Abstracts	37.8	13	54.1	13	56.8	13	6	8
Sociological Abstracts	28.2	10	46.5	11	53.5	11	3	2

TABLE 6.12 HIT RATE BY DATASET

Correlation between the hit rates and the redefinition indices was tested using Kendall's *tau b* correlation coefficient. The results are shown in table 6.13.

	0 hits	*<6 hits*	*<11 hits*
distribution (raw)	-0.039 (.855)	-0.026 (.903)	0.026 (.903)
distribution (refined)	0.013 (.951)	-0.077 (.714)	-0.077 (.714)

TABLE 6.13 CORRELATION BETWEEN HIT RATE AND REDEFINITION (FIRST ANALYSIS)

The *p* values, shown in parentheses above, are two-tailed. They indicate no evidence whatsoever of a relationship between hit rate and redefinition. In fact, by direct observation of table 6.12, we can deduce that no relationship between hit rate and redefinition is in evidence.

Another examination of the possible factor of hit rate was performed using a sample from the Medline dataset, in which the initial queries with hit rates of ten or fewer were identified. The number of hits for each query was compared with their numbers of modified queries, according to raw and refined index definitions. Correlation was tested using Kendall's *tau b* coefficient, with the results detailed in table 6.14 below.

TABLE 6.14 CORRELATION BETWEEN HIT RATE AND REDEFINITION (SECOND ANALYSIS)

Hits	Redefinition (raw)	Redefinition (refined)
6	0	0
0	1	1
0	2	2
1	0	0
1	5	4
1	0	0
0	5	1
3	1	1
0	23	22
0	1	0
7	7	7
9	1	0

0	11	10
10	1	0
0	1	0
6	6	2
0	3	2
2	3	2
8	8	8
0	13	13
5	1	1
4	5	5
tau b	-0.096	-0.113
p	0.569	0.508

The above *p* values indicate, again, no significant correlation between low hit rate and redefinition.

6.6 Retrievability

The levels of redefinition (in raw and refined index terms) occurring immediately following *kw* and *su* queries in the WorldCat dataset were compared and the proportions tested for association using Fisher's Exact test, as per table 6.15 below.

	raw		refined	
	redefined	**not redefined**	**redefined**	**not redefined**
kw searches	164	214	88	290
su searches	14	14	8	20
Fisher's test	*p*=0.5560 (two-tailed)		*p*=0.4965 (two-tailed)	

TABLE 6.15 COMPARISON BETWEEN *KW* AND *SU* SEARCHES

Given the *p* values above, the null hypothesis of non-association was not rejected. If there is a relationship between the two variables, it would appear to be weak. That is, the difference in retrievability does not appear to much affect redefinition.

Moreover, it is the *su* queries, rather than the *kw* queries, which appear to generate more redefinition, so if this were significantly more, it would invert the theory of retrievability engendering redefinition.

6.7 Discipline Area

Table 6.16 shows the specific discipline areas of the six databases aimed at particular audiences, as per table 5.14, and their respective rankings according to the refined index of redefinition.

	discipline area (DDC)	*redefinition (refined)*
ABI/INFORM	330	3
Business Management	330	1
Education Abstracts	370	6
Health & Wellness	610	4
Medline	610	5
Sociological Abstracts	300	2

TABLE 6.16 DISCIPLINE AREAS AND REDEFINITION

A Kruskal-Wallis test was performed to detect any association between discipline and redefinition, after Education Abstracts and Sociological Abstracts were grouped together given their similar disciplinary nature. The resulting H-statistic was 4.33 (p=0.115), which meant that no particular relationship between the discipline area and redefinition was detected. However, the low p value and direct observation suggested some association. Thus discipline *might* be a factor, even if a weak one. It might perhaps be enough of a factor to work against any correlation between abstract length and redefinition, for instance.

Chapter 7 CONCLUSIONS

7.0 Introduction

Let us now address directly the ten hypotheses that this thesis set out to investigate (see chapter 2), according to the evidence produced from the main analysis and pilot study.

7.1 Significance of Redefinition

(1) Information seekers redefine their information goals as a result of interaction with some bibliographic information retrieval systems significantly frequently.

In the pilot study, redefinition was identified in 37% of search sessions, a good deal higher than our benchmark of ten per cent for "significant" redefinition as a whole, proposed in section 2.11.

Although the indices developed for the main analysis are not intended to measure the *amount* of redefinition, or even strong redefinition, *per se,* it is nevertheless interesting to compare the estimated amount of redefinition taking place at the OPAC in the pilot study with an estimation of the amount of strong redefinition "picked up" by the refined index. This index is based on transaction logs in which the author identifies types of query modification that *may well* be caused by redefinition. One must bear in mind that some of the queries registered by the index were probably not the product of redefinition, but instead the product of a modification of search strategy. On the other hand, some modified queries not registered by the index probably were products of redefinition. Using the figures based on the coding of the WorldCat queries, we estimate that 26.3% (104/395) of search sessions, based on an initial query, included strong redefinition.

The main analysis dealt only with strong redefinition – that kind of redefinition acted upon immediately within the same search session. In the pilot study we have already seen how "weak redefinition" is also likely to occur, and quite

possibly to an even greater extent. Such redefinition does not lead to an immediate follow-up query on the same system, but might well lead to a follow-up search on a different system, or at least to a different selection of documents from which would otherwise have occurred. It is thus also of relevance to the systems designer.

7.2 Redefinition Index

(2) An index can be produced that accurately represents the relative frequencies at which information seekers redefine their information goals as a result of interaction with various bibliographic information retrieval systems.

As already pointed out, the indices employed in the main analysis are not measures of the amount of redefinition taking place *per se;* rather they are intended as reasonably reliable measures of the *relative* amounts of strong redefinition occurring across databases, which for our purposes equate to different information systems.

The indices are blunted, it is assumed, by a good deal of *noise.* That is, some of the queries registered in the indices, especially in the raw index, are actually not the product of redefinition, but instead the product of a modification of search *strategy.* Conversely, the refined index in particular is likely to have "missed" queries that were in fact the product of redefinition (at least in part). In these ways, the indices may show *smaller* differences between amounts of strong redefinition than would in fact be the case. Given this bluntness, if a statistical test on the indices indicates a "real" difference, we should thus be all the more confident in concluding that a real difference is indeed evident.

The two tests performed on the indices for control samples show no statistically significant difference (see section 6.2). If we assume that the nature of search behaviour would be similar from one day to the next, on the same database, then this result provides evidence that the indices do not produce false positives. That is, they do not indicate differences in search behaviour that are not there.

On the other hand, some of the tests performed on the indices for the datasets pertaining to the searching on the different databases do show statistically significant differences. This does not in itself mean that the indices accurately represent differences in the amounts of search redefinition, but query modification, on which they are based, can only otherwise be explained in terms of changing search strategy. The key assumption here is that strategy plays a reasonably uniform role in query modification across databases. If the statistically significant differences cannot be put down to differences of strategic behaviour, then they must be attributed to different amounts of redefinition. As such, the indices can be accepted as indices of redefinition.

Given this conclusion, we can then compare the two indices of redefinition. We find that the refined index consistently shows greater differences between redefinition than does the raw index. This consistency suggests that the refined index is a "sharper" one, as intended. In other words, the refined index is a more precise index than the raw index. It is thus recommended that the refined index would be a better candidate for further development.

7.3 Impact of System

(3) Information seekers redefine their information goals as a result of interaction with some bibliographic information retrieval systems more frequently than they do with others.

The Kolmogorov-Smirnov tests show statistically significant differences in the redefinition indices across various FirstSearch databases. For the purposes of this study, different FirstSearch databases represent different IR systems (they can be interrogated independently). Assuming that other aspects of users' search behaviour is uniform in at least some of the cross-dataset comparisons, we conclude that the nature of the information retrieval system can have an impact on redefinition.

The Kolmogorov-Smirnov tests show a distributional difference in the indices. Examining the shapes and central tendencies, we can observe that *on the average*

(however this might be defined) some systems produce more search redefinition than others.

7.4 Impact of Metadata

(4) Differences in the nature of the displayed content of bibliographic information retrieval systems affect the frequency at which information seekers redefine their information goals as a result of interaction with the systems.

In the pilot study, encountered metadata was attributed as the cause of 24% out of the 37% instances of redefinition. In the main analysis, comparison of indices of ArticleFirst and other full-text databases, indicated a difference which was probably due to ArticleFirst's lack of metadata (section 6.2), and the lack of evidence to support a "retrievability" factor (section 6.5) would point to the lack of *displayed* metadata being a factor working against search redefinition.

However, other factors appear to be at least as important, if not more so. Comparison of the indices for the full-text and non-full-text datasets indicated that redefinition on full-text databases occurs more frequently than it does on bibliographic-only databases (see section 6.3); indeed, the non-full-text average is clearly lower than any of those for the full-text datasets. This suggests that the presence of full-text is more of an effect on redefinition than are other differences between databases, including metadata differences. It might be that searching is more "ambitious" when the user is presented with the possibility of full text, which means more critical evaluation of results and review of the information goal.

Further, within the range of non-full-text databases, WorldCat would appear to induce significantly less redefinition than others, as comparison between WorldCat and other non-full-text databases showed (section 6.4). This could well be due to the difference of document representation, that is, WorldCat represents whole items (such as monographs), the other databases represent articles. Perhaps searchers are less "ambitious" on WorldCat due to additional effort

required to follow up their searches – they need to obtain and read larger documents, in general.

7.5 Abstract Presence

(5) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems which display abstracts more frequently than they do with bibliographic information retrieval systems which do not display abstracts (all other aspects being equal).

Comparison of indices for the Group Z dataset, representing databases without abstracts, and the full-text databases as a whole, showed that the presence of abstracts makes for more redefinition (section 6.2.1). Again, with the results indicating that retrievability may not be a significant factor, abstract *displays* would appear to encourage search redefinition. However, more research needs to be done on the retrievability of abstracts specifically, before we can firmly conclude that it is the displayed abstract, rather than the indexed abstract, that is the major factor.

Indeed, the correlation test on abstract frequency and redefinition index does not demonstrate the former to be a major cause of search redefinition. This does not imply that there is no relationship, and the relationship might be much stronger if users had given themselves more opportunity to display abstracts by opening up more full records. A low p value in the case of the refined index points to the need for further investigation. The number of times a searcher encounters abstracts might not necessarily be a large factor, but a factor nevertheless. If the difference in frequency is very wide (such as "routinely" versus never), then it may well become significant.

In any case, it seems likely that the proportion of records with abstracts is not *lineally* related to redefinition; there may be a logarithmic relationship, for example, such that the impact of abstract frequency on redefinition increases at a slower rate than does abstract frequency itself.

7.6 Abstract Length

(6) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems which display longer abstracts more frequently than they do with bibliographic information retrieval systems which display shorter abstracts (all other aspects being equal).

The correlation test on abstract length and redefinition index (see section 6.2.2) produced no evidence that redefinition is much affected by abstract length; indeed, the *p* values suggest that any factor based on abstract length would be very hard to detect. It may be that a causal relationship between redefinition and abstract length exists, but is very minor. In any case, it seems reasonable to suppose that the chances of redefinition resulting from the reading of an abstract do not increase much after a certain length.

7.7 Descriptors

(7) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems which display descriptors more frequently than they do with bibliographic information retrieval systems which do not display descriptors (all other aspects being equal).

Comparison of the indices for the Group Z and ArticleFirst datasets (section 6.2.3) showed a difference in the amount of redefinition which is probably due to the lack of descriptors and/or identifiers in the ArticleFirst database. (None of the databases includes abstracts.) In fact, the ArticleFirst indices are easily the lowest of all the full-text indices, so the lack of these metadata elements would appear to be a considerably important factor.

It could well be that descriptors and identifiers provide cues for search goal revision, but it is speculated that it is the *hyperlinked* nature of the descriptors in the FirstSearch system which is the largest factor here, especially since abstract

frequency does not appear to be a huge factor (abstracts are not hyperlinked). Searchers following the hyperlinks would log new queries that would have often registered in the redefinition indices. Furthermore, the ease with which hyperlinks can produce fresh result sets probably encourages more exploration of the databases, and this in turn is likely to increase search redefinition. Investigation of the "hyperlinking" factor is needed. For one thing, it is unclear how much the hyperlinks are utilized – which queries are so derived is not available on the FirstSearch logs. In the Excite context, evidence suggests that "related item" options are under-utilized (Spink, Jansen & Ozmultu 2000). If hyperlinking does turn out to be a large factor, we should bear in mind that this interface feature is not so much dependent on record content – it could be introduced for other metadata elements, including abstract keywords, for example.

7.8 Hit Rate

(8) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems more frequently when such interaction produces fewer relevant documents (in the view of the seeker).

The results of the analyses described in section 6.5 do not corroborate the conclusions drawn from observations made during the pilot study (chapter 4), where a significant amount of redefinition appeared to be caused by "search failure." The analysis produced no evidence at all that hit rate is a factor in the FirstSearch context, either positive or negative. It could be that search failure or a low hit rate is more of a factor on some systems than on others; that on public library OPACs, for instance, users tend not to think about their search so much before they begin typing in their initial query, whereas users of certain other systems (including FirstSearch) think more carefully about what it is they want (and how to get it), as they are going to greater trouble and expense. Or it could just be that the researcher's *perception* of frustration leading to redefinition, from the pilot study, was not borne out by the facts of the main analysis. One suspects

there might be an element of both involved, but further research is needed to shed more light on this question.

7.9 Retrievability

(9) Levels of precision and recall produced by a bibliographic information retrieval system affect the frequency at which its users redefine their information goals.

The results in section 6.6 do not show any evidence that greater retrievability leads to more redefinition, indeed, if there is any relationship between retrievability and redefinition, this analysis suggests that more retrievability makes for less redefinition, which seems counter-intuitive, as retrievability reduces precision, in general, and one would have thought low precision, like low hit rates, would encourage search goal review and redefinition.

On reflection, it may be that more "sophisticated" users who search on the subject fields as opposed to all the keyword fields, "care" more about their search and are more prepared to scrutinise results and follow up on redefinition. This would perhaps counter redefinition caused by less precision. It may be that a certain mid-level of precision is optimal for redefinition, with the user encountering citations which are "partially relevant," neither too relevant nor too irrelevant – this would tie in with the research into relevance carried out by Spink and Greisdorf (1997). But it may also be that notwithstanding this optimal level, precision is not a particularly important factor, and that no measure of relevance (of which there are many, of course) would necessarily be a good indicator of redefinition potential.

It is also important to bear in mind that retrievability is only one possible cause of less precision and more recall, and that the findings here do not exclude the chance of other precision/recall factors working against it, although the same dataset was used for both types of query, and the sample sizes were quite large.

7.10 Discipline

(10) Differences in the disciplines covered by different bibliographic information systems affect the frequency with which their users redefine their information goals.

The results in section 6.7 do not lead to the acceptance of this hypothesis, but neither can we conclude that there is no relationship between discipline area and redefinition. The *p* value is by no means high. It is very possible that discipline is a factor, and as such might reduce association between metadata content and redefinition. Data from searches on more databases in various discipline areas would help us decide whether discipline is a significant factor, and further research is most definitely called for.

7.11 Summary

We can summarise our conclusions with respect to the research question as direct responses to the ten hypotheses, as below.

(1) Information seekers redefine their information goals as a result of interaction with some bibliographic information retrieval systems significantly frequently.
Accepted, in terms of perceived significance for systems designers.

(2) An index can be produced that accurately represents the relative frequencies at which information seekers redefine their information goals as a result of interaction with various bibliographic information retrieval systems.
Accepted, though further tests needed to explore wider cross-system applicability.

(3) Information seekers redefine their information goals as a result of interaction with some bibliographic information retrieval systems more

frequently than they do with others.

Accepted, with respect to different bibliographic databases.

(4) Differences in the nature of the displayed content of bibliographic information retrieval systems affect the frequency at which information seekers redefine their information goals as a result of interaction with the systems.

Accepted, although these may not necessarily be the most important factors.

(5) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems which display abstracts more frequently than they do with bibliographic information retrieval systems which do not display abstracts (all other aspects being equal).

Accepted, but databases with some abstracts have not been shown to produce more redefinition than databases with many abstracts.

(6) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems which display longer abstracts more frequently than they do with bibliographic information retrieval systems which display shorter abstracts (all other aspects being equal).

No evidence was produced to accept this hypothesis; the analysis indicates that if any impact does exist, it is probably small.

(7) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems which display descriptors more frequently than they do with bibliographic information retrieval systems which do not display descriptors (all other aspects being equal).

Accepted in the case of descriptors with hyperlinking functionality.

(8) Information seekers redefine their information goals as a result of interaction with bibliographic information retrieval systems more frequently when such interaction reveals fewer relevant documents (in the view of the seeker).

No evidence was produced to support this hypothesis; the results suggested that hit rate is not much of a factor either way.

(9) Levels of precision and recall produced by a bibliographic information retrieval system affect the frequency at which its users redefine their information goals.

No evidence was produced to support this hypothesis, but neither can it be rejected.

(10) Differences in the disciplines covered by different bibliographic information systems affect the frequency with which their users redefine their information goals.

The evidence does not particularly support this hypothesis, but further investigation is needed.

8. Implications for Information Retrieval

8.1 Implications for System Design

8.1.1 Facilitation of Redefinition as a Positive Design Attribute

It is assumed in this study that one should measure the performance of an IR system in terms of the information needs or wants for which it was interrogated by the user. Although information seekers might benefit from information furnished them via a search prior to redefinition, in a way that in the *long term* is of more value to them than is the information furnished them via a search post redefinition, we are making no judgement about *effect* of information on users. Equally, we assume that the systems designer is unconcerned about the moral nature of information (for "good," for "bad") that an IR system provides its user.

Given these assumptions, search goal redefinition might be regarded as a positive phenomenon and that an IR system generating such redefinition regarded as superior to a system that does not. However, we can only come to this conclusion if given the same amount of vagueness of the user's initial query. If, on the other hand, one system can facilitate greater accuracy in the user's initial query, more than another system can, then this may negate the favourability of a system which produces more search redefinition. As Katzer and Snyder (1990) show, a better reference interview may very effectively tighten a user's conceptualisation of their information need; such a reference interview could now be undertaken by a system as well as by a human.

So search goal redefinition may be considered a positive measure of a system, given a certain level of initial search goal clarity. However, there is a further point to recognise before taking this stance. One must consider any *costs* involved – to the user and to the system designer – in redefinition. If there are significant costs, then one must weigh these against the *extent* to which redefinition increases the fulfillment of the users' information needs and desires.

We noted previously how browsing is an activity associated with "serendipity" and that information seekers often employ this mode of information seeking when they possess only a vague notion of their information needs or desires – at least in relation to a specific information resource – and require such a notion to be sharpened (that is, brought into focus). Browsing has taken on a whole new dimension with the advent of the Internet, on which users "surf." The point about surfing, in relation to this discussion, is that surfers can sometimes *lose control* of their information seeking activity. As their information "net" grows exponentially, surfers may redefine their goals to such an extent, that they become over-ambitious. They form in their mind's eye a vision of a record-breaking fish, and explore the depths of the ocean to see if such a fish might be caught, and in their enthusiasm let a perfectly good catch slip through their hands, one that would at least have satisfied their hunger.

Even if the IR system ensures that a user can "catch" all the information they encounter during their searching, that they decide is of utility, the user may still incur additional costs through redefinition. Primarily, there is the additional time involved – the user's time and the system's time. The latter may translate into telecommunication costs, for example, and possibly impact on other potential users. If simultaneous use of a system is limited, then the designer's objective needs to be cast in terms of "users" collectively, rather than "the user."

While we might consider that for some IR systems, additional system time is not a significant factor, the user's time must generally be considered significant to the user. However, it is exceptionally difficult to balance additional user's time against a user's additional information gain potentially resulting from redefinition. Such a balance depends on the importance of the information compared with the importance of the results of any activities that the user may otherwise carry out, during the time taken. We have already noted that we are to make no judgement about the impact of information on the user. Instead, what we must assume is that if a user redefines a search goal and continues to search using the IR system to meet the redefined goal, then it is worth the user's time to do so – that the user knows best.

We have found that in this study, users do sometimes wish to pursue their redefined search goal, and so we conclude that facilitation of redefinition is a positive system attribute. An IR system should allow the user to redefine their search goal whenever it *can* be redefined, should they wish to.

But should the system designer make a point of facilitating redefinition? *Could* the information gain resulting from redefinition be frequent and great enough for system designers to consider how redefinition might be encouraged in an inexpensive way? It is contended in this study that in some systems at least, redefined goals have been pursued on many occasions, perhaps representing a third of all searching, which suggests that even if not a main objective, systems designers should at least give the facilitation of redefinition considerable thought.

Some systems may make it easier for the user to pursue their redefined goal than do other systems. Of course, a system designer must strive to make it as easy as possible for the user to pursue their redefined goal – that is, minimise costs.

Given the above contentions, we shall discuss in the remainder of this chapter ways in which two objectives might be realized. The two objectives are:

(a) to facilitate search goal redefinition as much as possible

(b) to make the pursuit of redefined search goals (in the same system) as easy as possible.

8.1.2 Facilitation of Search Goal Redefinition

We concluded in the previous chapter that abstract presence, if not frequency, facilitated redefinition (at least through the FirstSearch interface). We also noted that it does not appear that indexing *necessarily* affects redefinition, in terms of its effect on precision and recall. We shall not explore this latter issue any further, suffice to say that more research needs to be undertaken to ascertain in what ways, if any, might indexing affect redefinition. Instead, we shall concentrate here on *display* issues.

This thesis contends that when record displays include abstracts, there is significantly more chance of the user redefining their search goals than when record displays do not include abstracts, all other things being equal. It would thus seem reasonable to call for abstracts to be included in all bibliographic records. However, there are costs involved in this, and we have also noted that some abstracts may be almost as effective as many abstracts. Abstracts appear to be a factor, but by no means the only one, and this has to be borne in mind when considering the costs involved in producing abstracts. It may be that to cover *all* citations with abstracts is not cost-effective, and that if abstracts can be obtained for a considerable proportion of citations, this is sufficient.

There are two obvious costs relating to abstracts. As well as the cost of creating or procuring the abstracts, there is the cost of the user's time in reading them (or parts of them). There seems no way in which these costs can be measured against the benefit of redefinition. We know that often the user may wish to pursue a redefined goal, but we do not know if, in hindsight, they would have done so given the amount of time it took to arrive at the redefined goal, through reading abstracts – they may well have read the abstracts for other reasons (such as to ascertain relevance in relation to the former search goal).

What we can propose is that bibliographic system designers should consider including abstracts for their records if they wish to provide a quality service and can obtain the abstracts without significant delay or overly burdensome financial outlay. We can also advise that abstracts need not necessarily appear in all records, nor do they need to be very long, for the purposes of inducing redefinition. However, the extent to which this is the case is likely to depend on the frequency with which abstracts are encountered through the system – if they are included at the immediate citation level, or the full records are displayed frequently by users (more frequently than by FirstSearch users), then this advice might not be so sound, and system designers should consider abstract frequency and length more carefully.

The nature and authorship of the abstract has not been investigated in this study. It is speculated that any abstract of reasonable quality (in traditionally defined

terms) may lead to significant amounts of redefinition, in which case automatically derived abstracts and those written by the authors of the texts would be worth including, rather than none at all.

It may well be that other bibliographic aspects of content, apart from the presence of abstracts, affect redefinition. Such aspects may in some cases be readily obtained, and thus worthy of serious consideration if only in terms of the benefits accrued through more redefinition. Some of the more prominent metadata elements (prominent in the user's mind and in presentation) may be particularly worth examining. Perhaps the descriptiveness or style of titles makes a big difference, or the presence of subtitles? We need to look in particular at any "cues" (to use Toms' term, 1997, 2002) that signpost the scanning of citation lists. The field of linguistics may be able to help here. In any case, further research in this area is clearly necessary.

As well the nature of bibliographic content, another possible factor that needs be investigated is the *presentation* of content. It may well be that the coverage of content on each screen is a factor on redefinition. If more of a record displayed with the initial citation makes for more redefinition, then the system designer needs to consider whether this would have any other effects. More of a record probably means fewer records on the initial screen. If a database contains few "relevant" materials, then this may not matter; but if it contains many, and if the user is impatient or lacks ambition, then it may well matter. Apart from quantity of record content, there is the area of presentation *style.* It may cost little to rearrange the elements of a record on a screen; only convention may dictate a particular order, and the user might not be especially upset if this convention is contravened. In addition, a system might be able to highlight "related concepts" for the user's consideration – where to position these on the screen without creating too much potential noise? Such questions require careful consideration and further research.

We should not be restricted in our investigations to traditional document representation either. We have seen in section 3.4 how some document retrieval systems have moved away from citation lists, and presented the user with more

visual displays. The aim has been to present a collection of documents as meaningfully as possible with respect to their subjects and content. This does not necessarily coincide with the goal of facilitating redefinition. Nevertheless, different visualisations appear to have some effect on selection and browsing behaviour, and it is very possible that they would also have an effect on the production of redefinition. It is hard enough to compare the effectiveness of different visualisation with regard to their main aim, as sated above, let alone their impact on a secondary variable such as redefinition. Yet this is research, perhaps with the help of psychology, that at some point should be attempted.

8.1.3 Ease of Redefined Searching

What became clear though both the pilot study and the main analysis was that the advent of hyperlinks has very likely had a large impact on the amount of *follow-up* on redefinition that nowadays takes place on document retrieval systems, if not on the amount of redefinition *per se.* A user realizes that a narrower term would more accurately represent their search goal; instead of having to start over and type in the new term, they simply click on the highlighted term within a citation and new search results appear. System designers have rightly observed that even two or three additional mouse and/or keyboard actions, or two or three additional screen redraws, can deter the user from implementing a new query.

The pilot study showed how some users took advantage of hyperlinked subject headings, while the results in the main analysis indicated that descriptors were a more important factor than abstracts with respect to redefinition. It seems likely that this is because the descriptors are hyperlinked (to other records containing the same descriptors), while the abstract words are not.

More hyperlinks, from more words in the record display, might thus be called for, in order to increase still further the amount of follow-up on redefinition. Not all contemporary systems present hyperlinks; some systems do not present hyperlinks on descriptors; and many systems do not present hyperlinks on words in titles or subtitles, or abstracts. It is perfectly feasible for a title or abstract to be initially displayed without individual words highlighted as hyperlinks, but for

them to appear as a cursor is moved over them. This would minimise distraction while giving the user a great deal more hyperlinking scope. A "synthetic hyperlinking" feature would also be attractive to more sophisticated users (as was noted during observations in the pilot study), whereby various words or phrases in a display, could be combined (in Boolean fashion), perhaps through right clicks, and then searched on. Admittedly, fairly sophisticated programming would be required to achieve this.

8.2 Implications for User Education

In this chapter, we have emphasised the role of systems design, but must not forget the importance of user education, which may done through the system or independently of it. Indeed, we should remember that redefinition itself is ultimately up to the user, and that they can play a very active role in honing their search goal concept. They can be educated to conceptualise a more accurate "mental model" not only of the system, but of the search process itself, and thus appreciate the difference greater thought about their information wants can make to their information gains. Users can also be taught to review their information goals through action, for example, by taking the time to display more full records and to read the abstracts.

Quality user education can likewise make a large difference to the pursuit of redefined search goals. Users need to be taught (apart from how to accurately conceptualise their search goal in the first place) how to follow up on new thoughts about their search goals as efficiently and effectively as possible, which may entail learning about the use of short-cut search methods, reviewing search history, and also, perhaps most importantly, an appreciation of the way in which revised queries may yield much better results.

Such objectives for user education are not especially new, but they are certainly worth reinforcing in relation to a subject as subtle as search goal redefinition.

8.3 Implications for User Modeling

As discussed in section 3.1, several models have been introduced into information science which attempt to integrate the "hard" and "soft" strands of the discipline, and which emphasise the interactivity of user and system. However, it is contended here that there is still an element that would appear to be missing from these models.

This research has shown that goal redefinition can take place frequently, yet the models do not leave much room for such redefinition. Ingwersen's "mediator model" (1992), for example, is based on a "cognitive viewpoint" whereby the user's cognition of their information need is shaped by their "problem space," which in turn is shaped by their "actual state of knowledge." This means that a user's cognitive state – including their concept of information need – can change as the user interacts with a system. However, only the acquisition of *external* knowledge is portrayed as a cause of this.

It is postulated that a major, perhaps more common, cause of goal redefinition is not the acquisition of "new" knowledge, but "old" knowledge remembered, through cues encountered during interaction with a system. These cues bring the user's unconscious knowledge into consciousness, providing the user with a better understanding of their information want. One reason why goal redefinition during search sessions has rarely been considered by IR researchers may well be that they have failed to recognize the role that the "unconscious" part of the mind plays in the information acquisition process.

This new element – the "unconscious" mind – is defined here as the contents of the user's mind of which the user is unaware at a given moment in time. This may include a very large part of their memory. It should be made clear that this definition pertains to cognitive science; it is *not* derived from psychoanalysis, and the new model which the author is proposing is not in any way psychoanalytic. In order to make this point all the more clear, the term "nonconscious" will be used instead of "unconscious" for the remainder of this discussion; its definition is the same as that rendered above.

Ingwersen's "cognitive viewpoint" excludes the possibility that internal cognitive processes may change the user's concept of information need independently of their "actual state of knowledge." The problem is that it assumes that given a certain "actual state of knowledge," the user constructs a particular concept of the information need. But is an information need always clearly and consistently conceptualised by the user? The model fails to recognize that a considerable part of a user's information want may not be conceptualized, especially at the outset, but instead resides as part of the user's nonconscious. This "nonconscious" aspect of information retrieval behaviour is not factored into any of the other models of information retrieval with which the author is familiar (the most important of which are covered by Wilson 1999, and Järvelin and Wilson 2003). Rather, cognition is viewed as independent of the nonconscious, when in fact it is anything but. Cognitive science has, on many occasions, shown how cognition is in constant interaction with the nonconscious part of the mind. As nonconscious thoughts become conscious ones, so may come a review of an information seeker's intentions.

Users may be able to articulate what they can *consciously* conceptualise, but they may not always be able to consciously conceptualise the entirety of their information want, and this incompleteness is represented by a lack of full expression of their information goal, something with which reference librarians are apparently quite familiar. Thus redefinition of an information goal may not only be induced by exposure to new external knowledge, but also by exposure to *old* knowledge, previously internalized, but the relevance of which only dawns on the user after initial conceptualization of the information goal.

Another problem with the models that attempt to describe information retrieval, as opposed to information seeking, from the perspective of this study, is that they tend to ignore the phenomenon of "information encountering." We have seen that a significant cause of goal redefinition is the encountering of metadata displayed *following* searching. IR researchers have commonly focused on the search itself, at the expense of *selection.* However, it appears that the notion of relevance may quite often be reviewed during the selection of documents, as represented by citation lists and metadata.

A model of user-system interaction must take into account both "hunting" and "gathering" modes of information acquisition: searching on the one hand, and on the other, encountering, browsing, scanning and selection. Both types of activity drive the information acquisition process, even during interaction with an IR system. Searches lead to selection; but during selection, users may encounter bibliographic cues that trigger new searches.

If researchers acknowledge the importance of encountering and evaluation in user-system interaction, they may find it easier to embrace the concept of fluid information goals, and, ultimately, fluid information needs and wants. By recognizing that encountering can take place even at the level of the search session, the element of the nonconscious may be included in a new information retrieval model. Indeed, this model can thus become part of a wider information seeking model, underpinned by a continuum, with hunting and conscious, or purposeful, searching at the one end, and with gathering and encountering, at the other. As we move towards the "gathering" end, so there is more scope for the nonconscious to impact on information acquisition, by becoming conscious. But the important point is that information behaviour is not fixed at a particular point on the continuum. Because there is constant interplay between conscious and nonconscious parts of the mind, there is always potential for behaviour to move along the continuum, and very often this is exactly what occurs. Information retrieval is thus integrated into the larger picture of information acquisition, by recognizing the way in which purposeful searching can be combined, and often is, with scanning and encountering.

A new user model of information retrieval is thus proposed, one in which the information seeker interacts with an environment which includes particular IR systems, each capable of prompting the user to change both direction and strategy – and this may occur within a particular search session as well as outside of it. Unlike other "ecological" models, this model attempts to integrate purposeful information seeking with incidental information acquisition by incorporating the nonconscious element of the human mind. Sometimes people have quite a clear information goal, other times they have a less clear goal, sometimes they have no particular goal. At any given moment, they would also

have a large number of nonconscious "information interests," ready to be identified and acted upon should a powerful enough cue make their owners aware, or conscious, of them.

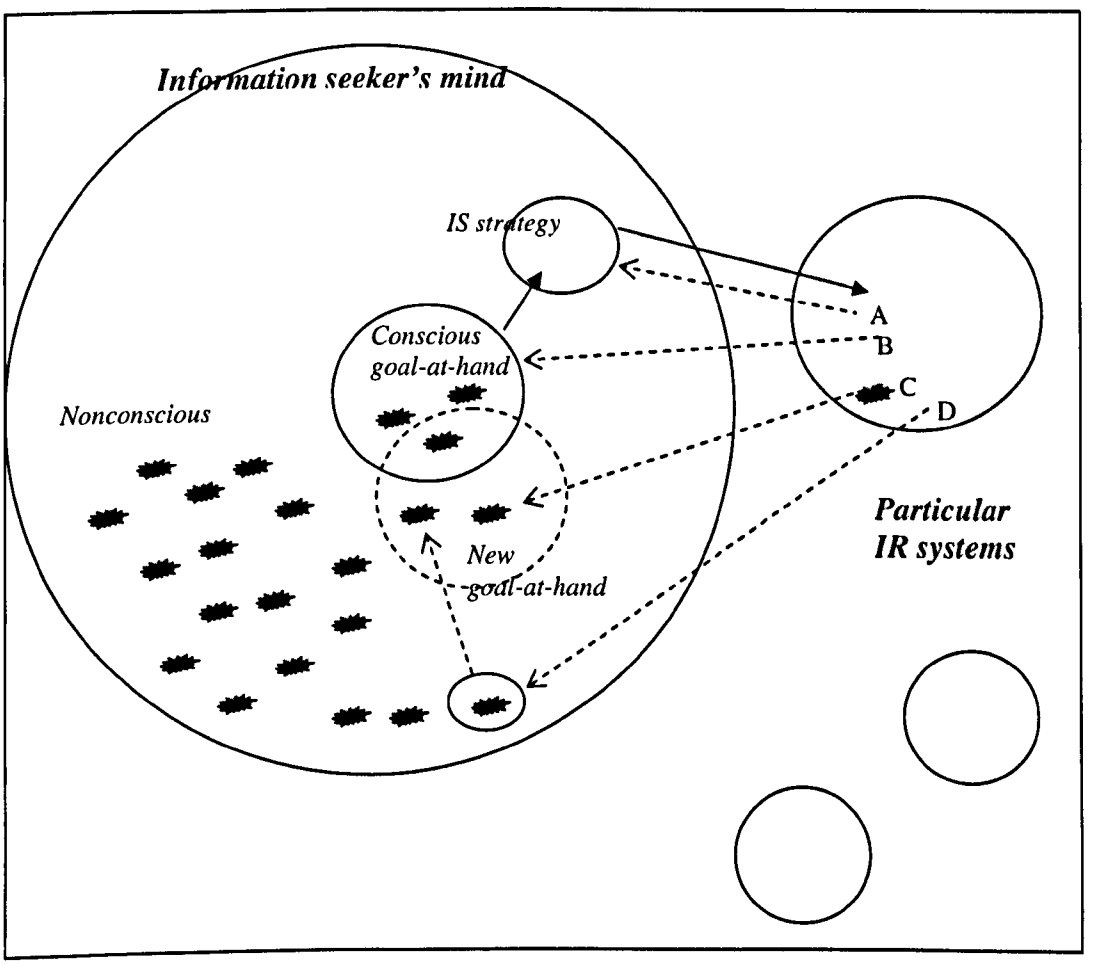
If an information seeker has a clear enough "goal-at-hand," that is, a goal to which they are applying themselves at a given moment, they may choose to use a particular IR system, such as a bibliographic database. Even then, however, the goal-at-hand may not be *crystal* clear. Previous studies have indicated how poorly formulated many real-life queries on OPACs and other IR systems are, not only poor in terms of search strategy, or because of a lack of knowledge of how to use the system, but also because of a lack of knowledge of what it is the user actually wants to find. What they cannot yet conceptualise, however, users may become conscious of later (whether or not they can express the concept in a way that will help is another matter).

Just as other unrelated information interests may surface in an information seeker's consciousness, given a particular cue, so might the user of a system become conscious of other *related* information wants, due to certain cues produced by the system's feedback (or from elsewhere). By *related,* what we mean here is that the information wants are relevant to the pursuit of the goal-at-hand.

Perhaps only a small part of a user's nonconscious could be related to a particular goal-at-hand, while there may be a much larger part ready to offer up any number of other goals for the information seeker to pursue instead of the goal-at-hand. However, many information interests – for example, that the document should be in a specific language – may be related, potentially, to more than one goal.

Figure 8.1 below summarises the author's model of information retrieval. The four broken arrows A-D are the four ways in which system feedback can influence information seeking. With A, ideas relating to search strategy are derived. With B, the information provided by the system meets the original information goal, or partially does so. With C, new information is taken on board which redefines the information goal. With D, cues from the system prompt the

user to redefine the information goal in terms of old knowledge that is moved into their consciousness.



= aspect of information interest

FIGURE 8.1 USER MODEL OF INFORMATION RETRIEVAL EMBRACING THE NONCONSCIOUS PART OF THE USER'S MIND AND INFORMATION ENCOUNTERING

This model is intended to cover the full range of information seeking behaviours, from the most specific searches, such as fact checking and known-item searches, where the information seeker is aware of most aspects of a search goal, to the most incidental, or even accidental, information encounter, which does not relate to any goal-at-hand (perhaps the actor is not even engaged in information seeking activities at the time). These behaviours are all dependent upon the changing make-up of the actor's cognition, not only in terms of new strategic opportunities that he or she becomes aware of, but also in terms of the new information wants

that make their way into the actor's consciousness and thus define or redefine his or her information goal-at-hand. Information seeking behaviours constantly evolve and metamorphise as the actor's cognitive state changes.

The model includes the possibility of information goal definition or redefinition through external cues which trigger awareness of the relevance of *old* knowledge, and also includes the possibility of information goal construction or transformation through the internalization of *new* knowledge encountered. New knowledge may not be internalized so frequently during interaction with a typical bibliographic retrieval system, and when encountered in day-to-day activities will often not be employed in an information seeking activity straightaway. Nevertheless, there may be times when system feedback throws up a new piece of information that can be used directly, either strategically, or to transform a search goal. For example, a user may *learn* from a record the scientific name for an animal – this information may be acted upon strategically. In another case, a user may *learn,* when searching for information on a chess opening, that there is a particular variation currently in vogue – the user may wish to refocus their search on this particular variation. This would represent a revision of the search goal.

However, although new knowledge may affect the direction of a search session, it may play a more important role in other information seeking contexts. Its role in user-system interaction may have been exaggerated, with *old* knowledge brought to consciousness being mistaken for "new knowledge." It is contended that, as well as new knowledge, the nonconscious would often come to the fore to redirect a search session, either in terms of strategy or goal, or both. For instance, a citation *reminds* a user of a synonym that they can then use in a query reformulation. Or after perusing citations, the chess player *realizes* that it is not just information about a particular opening that they require, but examples of the types of game that develop afterwards.

The author's model represents a view in which information encountering, leading to incidental information acquisition, is important in the traditional IR system context as well as in other, less directed information seeking contexts. Its

importance is based not only on what new information a system can provide, but on how system feedback can facilitate new states of consciousness. This study has not attempted to weigh the impact of new information against the potential of nonconscious knowledge, but instead advocates a model which incorporates the latter as well as the former, to help explain why information goal redefinition occurs as frequently as it does. This new model thus increases the scope for search goal redefinition.

9. Suggestions for Further Research

The literature review (chapter 3) revealed that little research has been done on search goal redefinition through interaction with information retrieval systems. While this study may have established that redefinition occurs significantly frequently on certain systems, other studies are needed to confirm or deny whether this is the case on other types of system, including non-bibliographic ones.

This study featured a chronological content analysis of an IR system's transaction logs which represents a new departure in information retrieval studies and TLA. The refined index used in the main analysis appears to reflect reasonably accurately relative frequencies of redefinition (see section 7.2), but it needs to be further validated using transaction logs from other systems and triangulated with other methodologies. As reported in chapters 4 and 5, other methodologies proved problematic for a range of reasons, some of them practical, others theoretical. The issue which caused greatest concern was that of obtrusivity. Given the nature of the subject, particularly as it has been brought out by the new user model proposed in section 8.3, involving the nonconscious, this issue is of particular importance. Obtrusive methods may lead to reflection that might not otherwise take place, and this in turn may produce more interplay between conscious and nonconscious.

Nevertheless, the methodological problems described in this study are not necessarily insurmountable, and a specially designed interface might produce valuable pre-, intra-, and post-session feedback from the user in a way that is reasonably unobtrusive. The challenge will be to design a research instrument that successfully combines a number of approaches, including transaction logging.

The conclusions reached in this study about the relationship between certain bibliographic features and redefinition also need to be confirmed and elaborated upon by other studies using other methodologies. It would appear that redefinition is not proportionately affected by abstract length, that the

relationship might rather be represented by a logarithmic function whereby after a few lines of abstract, the chances of redefinition tail off (see section 7.6). Is this in fact the case? Does it apply to other types of metadata, such as titles, and to metadata in general? It was found that the presence of descriptors increased redefinition. It seemed that a main reason for this would be their hyperlinking nature, which facilitates new searches (see section 7.7). This needs to be tested in further research.

It would be interesting to see if more subtle differences in record content might in fact make a difference given a "super-refined" index. It was recognized that even the "refined" index constructed for this study is quite "blunt" in that it relies heavily on interpretation. A more sophisticated coding scheme might be devised. While the reliability of the coding scheme used in the main analysis held up quite well, a more complex one might be better tackled by a computer. One would obviously need to consider the considerable programming effort involved, compared with the amount of time required to manually code the logs, which was manageable in this study, but might not be if a much greater quantity of data is to be utilised.

The indices used in this study were probably also blunted due to the fact that the bibliographic features being analysed were at the second level of display, in the FirstSearch system, and this might not have been reached in many of the search sessions, or rarely reached. If the variables were analysed more directly, then some of the results might have been "sharper." Such analysis would occur if, for instance, a system invariably displayed the features being analysed at the first level. More subtle factors on redefinition might thus be detectable.

The selection of variables for this study was constrained by the parameters of the FirstSearch logs, on which the main analysis was based. It would be interesting to look at a range of other content variables, besides those examined in this study, such as the nature of titles and subtitles, different subject vocabularies, and the availability/display of extracts from the text. The extent to which subject browsing functionality affects redefinition might also be researched, as could the impact of hyperlinks, apart from those from descriptors, on the follow-up of

redefinition. Many presentational variables would also be worth investigating, such as the order of bibliographic elements, relevance ranking and automatic search suggestions.

Moreover, the whole area of visualisation needs to be researched in relation to the phenomenon of redefinition. The ways in which the user's mind is affected by different styles of document representation is a rich area of study only just beginning to be tackled. Indeed, the style, location and timing of information presented to the user (be it bibliographic, full-text, suggestive, system-related help, or any other kind) has become a critical subject of study for all sorts of systems designers, and the field of HCI will no doubt continue to contribute findings (especially if more higher-level, task-oriented studies called for by Beaulieu (2000) are produced). IR researchers need to take heed of them, and add their own.

The relationships, if any, between search goal redefinition and search failure, the precision of result sets, and discipline areas all need further analysis (see sections 7.8-7.10). This study did not yield evidence to support any of these relationships, but this does not mean that such evidence does not exist. A more sophisticated view of search failure, not just in terms of hit rates, may be necessary; a more interactive methodology may be needed to shed light on the influence precision has on redefinition; and a greater representation of disciplines may offer researchers the chance to investigate the information seeking approaches of different user groups.

A distinction was made in the pilot study between "weak" and "strong" redefinition – what was meant by this was redefinition that led to an immediate follow-up on the system, versus no such follow-up. Yet weak redefinition could still entail the redirected investigation of result sets: a user could redefine their search goal on perusal of a citation and go on to judge other records in a result set according to a revised relevance criteria. And even if it does not, there still may be large variations in the amount of weak redefinition occurring. Are the factors the same as they are for strong redefinition? Quite possibly, other factors are involved. While some work on the way in which users scan documents and

their surrogates has been conducted, the way in which relevance criteria are revised *during* the selection of documents from results sets has not yet been specifically researched.

This study focused on topical shifts of search goal, which throws up the question of whether the same levels of and factors on redefinition are to be found in the case of other elements and types of information goal. Given the emphasis on topicality in the retrieval mechanisms of most real-life systems, other methodologies apart from TLA are likely to be needed to shed light on this question.

Many users now operate across systems in a way that was perhaps unimaginable just a decade or so ago. They can bring up other windows, and copy and paste queries, with a few mouse movements or keystrokes. Nowadays, we should not only be looking at the results of redefinition within one search session on one system, but how it might be followed up on other systems, or back on the same system, during the whole of a *workstation* session. The classic TLA problem of session demarcation, which was also encountered in this study, becomes all the more pertinent.

Researchers need to create new opportunities for the collection of data, as well as creating new hypotheses. They need to address the more practical problems as well as theoretical ones. Obtaining clearance to use the transaction logs of real-life systems is not always easy, due to their commercial nature. Often, there are also legal and ethical issues. One would like to see researchers with access to more of the logs of leading IR systems.

This particular study would have benefited not from more data, but from more *useable* data. In other words, transaction logs need to be more detailed. Ideally, screen recordings should be obtained in order for the whole picture of interaction to be seen. Recording software, such as that mentioned in section 5.1.5, could be run in front of real-life systems, and experimental systems which allow for controlled adjustment of variables with real-life content and real-life clientele. With such software and systems in place, researchers would have much richer

data with which to work, and be in a better position to examine deeper questions of user-system interaction.

A new user model was put forward in section 8.3 as an attempt to explain more clearly why we saw so much goal redefinition taking place during real-life interaction with a bibliographic retrieval system. This model now needs to be tested, using data derived from a variety of sources. It is a model which is bound to require further development, on the part of both information seeking and information retrieval sides of the discipline, but if it is one which proves to be substantially valid, then perhaps we would have at our disposal a platform on which to build a more collaborative and integrated disciplinary effort.

APPENDIX A

Search Commands on FirstSearch for Medline Database

Parameter	*Label*
Keyword	kw:
Abstract	ab:
Accession number	no:
Address	ad:
Article type phrase	at=
Author	au:
Author phrase	au=
Central concept	cc:
Chemical substance	cs:
Citation owner	ow:
Corporate author phrase	co=
Date of entry	da:
Date of last revision	ud:
Division	sb:
Grant information	gi:
Identifier	id:
Identifier phrase	id=
Issue	is:
Language phrase	ln=
Medline number	mi:
MeSH heading	mh:
MeSH heading phrase	mh=
Molecular sequence	ms:
Named person phrase	na=
Notes	nt:
Number references	nr:
Page (first)	pg:
Place of publication	pl:
Publication date	pd:
Record type phrase	rt=
Record status phrase	rs=
Registry number	rn:
Source	so:
Source phrase	so=
Standard number	sn:
SubHeading (MeSH)	sh:
SubHeading (MeSH) phrase	sh=
Subject	su:
Subject All	sa=
Subject phrase	su=
Title	ti:
Title phrase	ti=
Volume	vo:
Year	yr:

exact phrase	*quotes* " "
plurals	+
variants or part of a word	*, #, or ?
all words	AND
one or all words	OR
one word but not another	NOT
words near each other, given order	w
words near each other, any order	n

APPENDIX B

Extract from Coded Log (WorldCat Queries)

Session	*Hits*	*Query*	*Code*
233660	2	kw: guitar and kw: components	S
233660	5969	kw: guitar and kw: parts	MS
233660	2284	kw: guitar and kw: strings	MS
233660	5969	kw: parts and kw: guitar	X
233660	0	kw: guitar-instructino	MS
233660	0	kw: guitar-instruction	X
233660	5381	kw: guitar and kw: instruction	X
233691	17	kw: cardwell and kw: australia	S
233691	13	kw: cardwell and kw: australia and dt= "bks"	NT
233691	0	(kw: australia and kw: sail) and kw: wife	MS
233691	327	kw: australia and kw: wife	MS
233691	222	kw: australia and kw: wife and dt= "bks"	MC
233691	222	kw: australia and kw: wife and dt= "bks"	X
233691	220	ln= "eng" and ((kw: australia and kw: wife and dt= "bks"))	MC
233692	0	ti: nutmegs and ti: ginger and dt= "rec"	S
233692	2	ti: nutmeg and ti: ginger and dt= "rec"	X
233692	3	nt: nutmeg and nt: ginger and dt= "rec"	X
233692	2	nt: nutmigs and nt: ginger and dt= "rec"	X
233692	6	nt: nuttmigs and nt: ginger and dt= "rec"	X
233692	0	nt: nuttmig and nt: ginger and dt= "rec"	X
233692	0	ti: nuttmigs and ti: ginger and dt= "rec"	X
233692	0	nt: nuttmig and nt: ginger and dt= "rec"	X
233692	3	nt: nutmeg and nt: ginger and dt= "rec"	X

233692	5	nt: nutmegs and nt: ginger and dt= "rec"	X
233692	2	nt: nutmigs and nt: ginger and dt= "rec"	X
233692	6	nt: nuttmigs and nt: ginger and dt= "rec"	X
233692	0	nt: nuttmegs and nt: ginger and dt= "rec"	X
233692	0	nt: nuttmeg and nt: ginger and dt= "rec"	X
233692	5	nt: nutmegs and nt: ginger and dt= "rec"	X
233692	3	nt: nutmeg and nt: ginger and dt= "rec"	X
233692	5	nt: nutmegs and nt: ginger and dt= "rec"	X
233692	6	nt: nuttmigs and nt: ginger and dt= "rec"	X
233692	0	nt: nuttmig and nt: ginger and dt= "rec"	X
233700	0	au: Impure and au: Reason	K
233700	12	ti: Impure and ti: Reason	K
233700	8	au: Sadji, and au: Uta	K
233700	2	au: Harris-Schenz, and au: Beverly	K
233709	2	ti: Robert and ti: Copland and ti: poems	K
233709	1	(au= "Erler, Mary Carpenter,") and au= "1937-"	K
233709	2	ti: Robert and ti: Copland and ti: Poems	K
233714	8	(ti: true and ti: north and ti: memoir) and au: conway and ln= "eng" and dt= "bks"	K
233726	437	ti: sacred and ti: places	K
233726	0	au: izu, and au: ken	K
233726	11	au: izu, and au: kenro	K
233731	7	au: Ruskin and ti: stones and yr: 1875	K
233731	1	au: Ruskin and ti: stones and yr: 1874	K
233731	1	au: Ruskin and ti: stones and yr: 1876	K
233731	10	au: Ruskin and ti: stones and yr: 1886	K
233731	4	au: Ruskin and ti: stones and yr: 1910	K
233731	5	au: Ruskin and ti: stones and yr: 1906	K
233736	56	au: dafermos	K
233736	9	au: bressan, and au: alberto	K
233738	540	kw: alfred and kw: marshall	S

233738	136	(su= "Marshall, Alfred,") and su= "1842-1924."	X
233740	45	ti: stranger and ti: stop	K
233740	2	(ti: stranger and ti: stop) and (ti: cast and ti: eye)	K
233740	0	au: Jacobs, and au: C and au: Walker	K
233740	29	au: Hormann and au: Ludwig	K
233740	227	au: Labbe and au: Philippe	K
233740	3	(au: Labbe and au: Philippe) and ti: Thesaurus	K
233740	0	au: Svertius	K
233740	72	au: Sweerts	K
233740	0	au: Sweerts and au: Franciscus	K
233740	0	au: Sweerts and au: Frans	K
233740	3	ti: Epitaphia and ti: ioco-seria	K
233740	231	au: Fogle, and au: Bruce	K
233740	1	(au: Fogle, and au: Bruce) and ti: Interrelations	K
233740	103	au: Linzey	K
233740	39	au: Linzey and au: Andrew	K
233740	1	au: Kurz, and au: Gary	K
233740	3	au: Laird, and au: Albert and au: B	K
233740	18	au: Laird, and au: Albert	K
233740	4	ti: Latin and ti: verse and ti: inscriptions	K
233740	1	au: Lanci-Altomare	K
233740	1	au: Tiedt, and au: Ernst	K
233740	33	au: Suffling, and au: Ernest	K
233740	144	au: Richard, and au: Lucien	K
233740	1	(au: Richard, and au: Lucien) and kw: Annuaire	K
233755	5	ti: my and ti: argument with "the" and ti: gestapo	K
233765	12	kw: Frank and kw: Potter's and kw: Science and kw: Gems and dt= "url"	S
233766	156	kw: shoemaking	S
233766	3	au= "Haire, Jesse S."	K
233766	162	su= "Sand Creek Massacre, Colo., 1864."	S

233766	601	(su= "Boots and shoes") and (su= "Trade and manufacture.")	MS
233766	103	(su= "Boots and shoes") and su= "History."	MS
233768	2	ti: sailing and ti: pilots and ti: bristol and ti: channel	K
233769	11	au: klyza, and au: christopher and dt= "bks"	K
233769	129	au: gonzalez, and au: george and dt= "bks"	K
233778	4	ti: dana and ti: family and ti: america	K
233785	14	kw: foster and kw: martin and dt= "mix"	S
233785	16	kw: hayes and kw: frank and dt= "mix"	S
233785	194	kw: palmer and kw: alexander	S
233785	19	kw: palmer and kw: alexander and dt= "mix"	NT
233790	0	kw: dilantin and kw: fda and kw: history	S
233790	0	kw: dilantin and kw: fda	BS
233790	53	kw: dilantin	BS
233790	31	kw: dilantin and dt= "bks"	MC
233796	4	(kw: embracing and kw: jesus) and (kw: the and kw: goddess)	S
233796	0	kw: gods and kw: who and kw: walk and kw: amoung and kw: us	MS
233796	1	kw: gods and kw: who and kw: walk and kw: among and kw: us	X
233796	14	kw: whole and kw: duty and kw: man and kw: according	S
233796	20	kw: women and kw: afghanistan and kw: under and kw: taliban	S
233798	190	ti: annie and ti: oakley	K
233798	65	ti: annie and ti: oakley and dt= "vis"	K
233811	0	se: factory and (se: industrial and se: management)	K
233811	23	ti: factory and (ti: industrial and ti: management)	K
233811	19	ti: american and ti: industry and ti: europe	K
233822	56	ti: advice and ti: young and ti: man	K
233822	1	ti: advice and ti: young and ti: man and dt= "art"	K
233826	16	kw: fuss, and kw: adam	S
233826	16	kw: fuss, and kw: adam	X
233863	0	au: Watts, and au: Rikk	K
233863	32	au: Toews, and au: John and au: B.	K

233879	240	ti: modeling and (ti: role and ti: modeling)	K
233879	3	(ti: modeling and (ti: role and ti: modeling)) and au: erickson	K
233884	21	au: Idowu, and au: E. and au: Bolaji	K
233887	7	au: herbst, and au: adolf	K
233889	1	au: huegel and ((kw: successful and kw: praying)) and ln= "eng" and dt= "bks"	K
233889	195	su: automobile and su: repair and su: shops and ln= "eng" and dt= "bks"	S
233897	0	kw: Sarajevo and kw: Children and dt= "vis" and li: EDK	S
233897	4	kw: Sarajevo and li: EDK	BC
233897	0	ti: "The" and ti: Contender and li: EDK	K
233897	149	ti: "The" and ti: Contender	K
233897	18	ti: "The" and ti: Contender and dt= "vis"	K
233897	744	kw: Yugoslavia and dt= "vis"	MS
233897	58	kw: Yugoslavia and li: EDK	MS
233923	6067	kw: developmental and kw: education	S
233923	92	kw: developmental and kw: education and dt= "url"	NT
233923	128	kw: developmental and kw: education and dt= "ser"	NT
233923	1	(au= "University of Texas at Brownsville,") and (au= "Developmental Education Dept.")	K
233930	13	su: euthanasia and kw: methods	S
233930	3003	su= "Euthanasia."	BS
233930	80	(su= "suicide, assisted.") OR (su= "technology, medical.") and su= "Euthanasia."	MS
233933	14	ti: taking and ti: offensive	K
233933	11	ti: taking and ti: offensive and dt= "bks"	K
233933	3	(au= "MacGarrigle, George L.,") and au= "1930-"	K
233934	2615	kw: modern and kw: art and yr: 1950-1959	S
233937	0	kw: belgian and kw: business and kw: etiquette	S
233937	2086	kw: business and kw: etiquette	NS
233937	1401	kw: business and kw: etiquette and dt= "bks"	MC
233937	14	kw: european and kw: business and kw: etiquette	MS
233937	19	(su= "Business etiquette") and su= "Europe."	X
233937	19	(su= "Business etiquette") and su= "Europe."	X

233937	6	kw: belgium and kw: business and kw: etiquette	X
233937	147	kw: international and kw: business and kw: etiquette	MS
233937	97	kw: international and kw: business and kw: etiquette and dt= "bks"	MC
233937	14	kw: international and kw: business and kw: etiquette and dt= "url"	MC
233950	146	kw: letter and kw: emily and dt= "bks"	S
234016	1	kw: Lawrence and kw: Birken	S
234016	0	ti: Lawrence and ti: Birken	K
234016	6	au: Lawrence and au: Birken	K
234016	23	ti: Baechler	K
234016	1	ti: Capitalism and ((ti: the and ti: rise and ti: west))	K
234033	16	au: gordon, and au: maggi and au: mccormick	K
234048	2212	au: Jackson, and au: George and au: A.	K
234048	10	(au: Jackson, and au: George and au: A.) and kw: Preliminary	K
234048	13	kw: Stokes, and kw: I. and kw: N. and kw: Phelps	S
234048	2	kw: Hawes-Stokes and kw: Collection	MS
234048	489	kw: Austin, and kw: William and kw: D.	S
234048	158	au: Austin, and au: William and au: D.	K
234048	3	ti: History and ti: Boston and ti: Society and ti: Architects	K
234048	5	ti: photographers and ti: sourcebook	K
234051	1	kw: danas and ((kw: the and kw: dana and kw: farm))	S
234062	1	sn: 0789714868	K
234068	6	ti: last and ti: cape and ti: horners	K
234069	197	ti: kentucky and ti: war and ti: 1812	K
234069	0	kw: kentucky and kw: war and kw: 1812 and kw: quisenberry	S
234071	1031	kw: truth and kw: last	S
234071	715	kw: truth and kw: last and dt= "bks"	NT
234071	3	(ti: truth and ti: last) and au: keeler	K
234076	6	kw: worldwide and kw: offshore and kw: owners	S
234080	1083	au: meyer, and au: john and dt= "bks"	K
234085	4	au: Barbara and au: Hancin-Bhatt	K

234085	184	ti: Second and ti: Language and ti: Research	K
234085	10	ti: Second and ti: Language and ti: Research and dt= "ser"	K
234085	5	ti: Second and ti: Language and ti: Research and dt= "url"	K
234087	5	au: silberman and ((ti: hidden and ti: scrolls))	K
234095	2	(au: larsson and ti= "concertino") and ((kw: double and kw: bass))	K
234095	19	(su= "Concertos Double bass with string orchestra") and su= "Scores."	S
234106	225	kw: Chords, and kw: intervals	S
234106	1218	kw: Clough	K
234106	2411	au: Clough	X
234106	22	(ti: Chords, and ti: intervals,) and ti: scales	NS
234106	4	(ti: scales, and ti: intervals, and ti: keys) and ti: triads	MS
234106	4	ti: "the" and ti: great and ti: irish and ti: potato and ti: famine	K
234125	1879	kw: french and kw: revolution and kw: bibliography	S
234137	2	kw: expectancy and kw: theory and yr: 1990-2002 and ln= "eng" and li: IVU	S
234137	2	kw: expectancy and kw: theory and ln= "eng" and li: IVU	MT
234142	7	kw: audio and kw: engineering and kw: society and kw: journal and ln= "eng" and dt= "ser"	S
234142	0	(kw: audio and kw: engineering and kw: society and kw: journal) and kw: preprint and ln= "eng" and dt= "ser"	NT
234142	7	kw: audio and kw: engineering and kw: society and kw: journal and ln= "eng" and dt= "ser"	X
234142	174	au= "Audio Engineering Society."	K
234145	1	nb: 083302955x	K
234147	0	sn: 0763578533	K
234147	0	se: Sports and se: History and se: story and se: Baseball	K
234147	32	au: suen, and au: anastasia	K
234147	53	au: feldman, and au: heather	K
234147	0	sn: 0763578436	K
234147	4	ti: Wayne and ti: Gretzky and ti: Hockey and ti: Star	K
234147	0	ti: basektball and ((au: otten, and au: jack))	K
234164	1	(kw: crying and kw: woman) and ((au: brian and au: naze))	K
234165	7	kw: fungus and kw: link	S
234165	5	kw: fungus and kw: link and dt= "bks"	NT

234165	3	(au= "Kaufmann, Doug A.,") and au= "1949-"	K
234170	123	ti: unanswered and kw: question	S
234171	15	au: garcia-aguilera and dt= "bks"	K
234184	5	ti: amerasia and ti: journal	K
234184	21	ti: ethnic and (ti: racial and ti: studies)	K
234184	2	ti: ethnic and (ti: racial and ti: studies) and dt= "ser"	K
234186	12	au: garrett, and au: charles and yr: 2000-2002	K
234186	3	au: garrett, and au: charles and au: l and yr: 2000-2002	K
234198	1	ti: new and ti: hydrocarbon and ti: potential and ti: bighorn and ti: basin	K
234201	169	au: furlani	K
234221	0	au: Sundberg and su: Vibrato and yr: 1989	K
234221	235	kw: Sundberg	S
234221	0	kw: Sundberg and kw: vibrato	NS
234221	0	kw: Sundbert and kw: 1989	NT
234221	10	kw: Sundberg and kw: 1989	X
234221	327	kw: Bennett and kw: 1981	MC
234221	0	kw: Bennett and kw: 1981 and kw: vibrato	MC
234221	5	kw: Bennett and kw: 1981 and kw: voice	MC
234221	1	kw: vibrato and kw: bennett	MC
234221	1	kw: Bennett and kw: vibrato	X
234221	359	kw: Bennett and kw: voice	MC
234221	0	kw: Bennett and kw: Dejonckere	MS
234221	2	kw: Dejonckere	MS
234223	15	kw: tithe and kw: applotment	S
234223	152	kw: griffith's and kw: valuation and kw: ireland	MS
234232	12	au: Ann and au: Laura and au: Stoler	K
234232	8	au: Robert and au: Haskett	K
234233	3	(ti: Citizenship and ti: Rites) and au: feinman	K
234233	3	(ti: Constructing and ti: Black and ti: Masculine) and au: wallace	K
234233	3	(ti: Constructing and ti: Black and ti: Masculine) and au: wallace	K

234233	3	(ti: Constructing and ti: Black and ti: Masculine) and au: wallace	K
234234	18	kw: sanfermines and ln= "spa" and dt= "bks"	S
234234	35	su= "Fiesta de San Fermín, Pamplona, Spain."	BT
234234	32	(su= "Pamplona Spain") and (su= "Social life and customs.")	MS
234234	17	(su= "Pamplona Spain") and (su= "Social life and customs.") and dt= "bks"	MS
234234	14	(su= "Pamplona Spain") and (su= "Social life and customs.") and ln= "spa" and dt= "bks"	MS
234234	5	su= "Fiesta de San Fermín."	X
234252	14	kw: great and kw: minds and kw: western and kw: intellectual and kw: thought and ln= "eng" and dt= "rec"	S
234273	4	au: mitchell and kw: indecision	K
234273	33	(au= "Mitchell, John Kearsley,") and au= "1798-1858."	K
234289	1	sn: 9573237067	K
234289	1	sn: 9577086365	K
234289	1	sn: 9577085040	K
234304	13	ti: Pontiac and ti: Press	K
234332	1	(kw: sorority and kw: history) and ((kw: alpha and kw: sigma and kw: tau))	S
234332	71	(kw: fraternal and kw: organization) and kw: history	MS
234332	86	(kw: greek and kw: organization) and kw: history	MS
234332	0	(((kw: greek and kw: organization)) and kw: history) and co= "panhellenic"	K
234332	0	(pb: panhellenic and kw: history) and co= "utsa"	K
234332	0	(pb: sorority and kw: history) and co= "utsa"	K
234332	3	(kw: sorority and su: organizations) and kw: history	MS
234332	1	(su: sorority and su: organizations) and kw: history	X
234357	1	ti: information and ti: sources and ti: art and ti: design and yr: 2000-	K
234357	0	ti: warhol and au: finck and yr: 2000-	K
234357	1	sn: 3925520627 and yr: 2000-	K
234357	0	(ti: imaginares and ti: museum) and au: muller and yr: 2000-	K
234357	0	(ti: imaginares and ti: museum) and au: muller	K
234357	0	ti: picassos and au: muller	K
234357	1	sn: 3775711236	K
234357	0	ti: catalogue and ti: ibeji	K

234357	1	ti: impact and ti: soutine	K
234357	1	ti: face and ti: portraits and ti: five	K
234357	2	ti: masterpiece and ti: rediscovered and ti: woman	K
234357	0	ti: belle and ti: europe and ti: temps	K
234357	0	sn: 2930236167	K
234357	1	sn: 9074822428	K
234357	1	ti: malevichs and ti: circle	K
234357	0	ti: kiefer and ti: sieben	K
234357	0	sn: 3775711244	K
234357	0	sn: 260103274	K
234357	1	ti: sidaner and ti: jardin and ti: gerberoy	K
234357	1	ti: philippe and ti: starck and ti: notoriete	K
234357	3	ti: duban and ti: couleurs	K
234357	1	ti: temps and ti: caravagisme	K
234357	5	ti: rodin and ti: bourgeois and ti: calais	K
234357	2	ti: learning and ti: construction and ti: failures and ti: forensic	K
234357	2	ti: surrealism and ti: desire and ti: unbound	K
234357	1	ti: studies and ti: medieval and ti: glass and ti: monasticism	K
234357	14	ti: ambassadors and ti: secret	K
234357	1	ti: courtly and ti: gardens and ti: holland	K
234357	5	ti: fortuna and ti: all	K
234357	0	ti: fortuna and kw: folger	K
234357	10	ti: fortune and kw: folger	K
234357	13	ti: artists and ti: multiples	K
234357	0	ti: renoir and ti: ecrits	K
234357	0	ti: renoir and ti: ecrits	K
234357	1	ti: retable and ti: fra and ti: angelico	K
234357	10	ti: tall and ti: buildings and ti: urban and ti: habitat	K
234360	1	(ti: sleep, and ti: arousal) and ti: performance	K
234411	2	(ti: Modern and ti: Techniques) and (ti: Technology and ti: Mtt)	K

234411	1	(ti: science and ti: bought) and ti: sold	K
234462	0	kw: genetic w discrimination AND kw: workplace and li: IUP	S
234462	0	kw: genetic w discrimination AND kw: employees and li: IUP	MS
234462	2	kw: genetic w discrimination and li: IUP	BS
234468	0	kw: boss and kw: 302 and kw: chassis and kw: modification	S
234468	4	kw: mustang and kw: boss and kw: 302	MS
234468	0	kw: ford and kw: 289-302	MS
234468	44	kw: ford and kw: boss	MS
234471	5	au: moors and au: annelies	K
234485	3	au: pocs and kw: witches	K
234491	16	kw: Los and kw: Sanfermines	S
234491	13	kw: Los and kw: Sanfermines and dt= "bks"	NT
234491	167	su= "Spain" and (su= "In literature.")	MS
234491	32	(su= "Pamplona Spain") and (su= "Social life and customs.")	MS
234491	17	(su= "Pamplona Spain") and (su= "Social life and customs.") and dt= "bks"	MC
234494	0	kw: albee and ((kw: virgina and kw: woolf)) and (dt= "bks" or dt= "ser" or dt= "art" or dt= "mix" or dt= "url")	S
234494	77	kw: albee and kw: woolf and (dt= "bks" or dt= "ser" or dt= "art" or dt= "mix" or dt= "url")	BS
234494	68	kw: albee and kw: woolf and ln= "eng" and (dt= "bks" or dt= "ser" or dt= "art" or dt= "mix" or dt= "url")	MC
234494	1	kw: albee and kw: woolf and dt= "art" and ln= "eng" and (dt= "bks" or dt= "ser" or dt= "art" or dt= "mix" or dt= "url")	NT
234494	69	kw: albee and kw: who's and ln= "eng" and (dt= "bks" or dt= "ser" or dt= "art" or dt= "mix" or dt= "url")	MC
234494	1	kw: albee and kw: who's and dt= "art" and ln= "eng" and (dt= "bks" or dt= "ser" or dt= "art" or dt= "mix" or dt= "url")	NT
234494	2	kw: albee and kw: who's and dt= "mix" and ln= "eng" and (dt= "bks" or dt= "ser" or dt= "art" or dt= "mix" or dt= "url")	NT
234494	2	kw: albee and kw: who's and dt= "url" and ln= "eng" and (dt= "bks" or dt= "ser" or dt= "art" or dt= "mix" or dt= "url")	NT
234503	325	au: Caillois and au: Roger	K
234503	8	su= "poetry." and (au: Caillois and au: Roger)	K
234511	8	ti: "the" and ti: divine and ti: comedy and li: RCL and (dt= "bks" or dt= "art" or dt= "url")	K
234516	2	kw: geschichte and kw: des and kw: orgelspiels and kw: Beispielband	S
234531	3	kw: Alvan and kw: Clark and kw: Sons, and kw: Artists and kw: Optics	S

234531	1	(su= "Telescope makers") AND (su= "United States") and su= "Biography."	MS
234531	77	(su= "Astronomical instruments") and su= "Catalogs."	MS
234534	12	ti: several and ti: regiments and ti: carolina	K
234534	5	(ti: north and ti: carolina and ti: troops) and au: manarin	K
234540	0	ti: Internation and ti: Bridge and ti: Conference	K
234540	64	ti: International and ti: Bridge and ti: Conference	K
234540	1	(au= "International Bridge Conference") AND au= "2000" and (au= "Pittsburgh, Pa.")	K
234553	9	ti: Excavations and ti: Giza	K
234565	2961	su: gay and su: men and su: fiction	S
234565	2913	su: gay and su: men and su: fiction and dt= "bks"	NT
234565	375	li: OCP and (su: gay and su: men and su: fiction and dt= "bks")	X
234571	3	(ti: peel and ti: slowly) and ti: see	K
234571	104	au= "Velvet Underground Musical group"	K
234573	0	(kw: anonymous and kw: genetic and kw: test) and (kw: huntington and kw: disease)	S
234573	0	kw: anonymous and kw: genetic and kw: test	BS
234573	129	kw: huntington and kw: disease	BS
234594	0	au: Lagardier	K
234594	518	ti: Geschichte and ti: von and ti: Stadt	K
234594	8	ti: Geschichte and ti: von and ti: Stadt and ti: und and ti: Landschaft	K
234607	1	ti: gaia's and ti: garden and yr: 2001 and dt= "bks"	K
234615	19	kw: joanne and kw: akalaitis	S
234642	203	kw: potato and kw: processing and dt= "bks"	S
234642	0	(kw: potato and kw: processing) and kw: dryers and dt= "bks"	NS
234647	0	ti: candlelighting and ti: encyclopedia	K
234647	1	ti: candle and ti: lighting and ti: encyclopedia	K
234647	1	au: tina and au: ketch	K
234650	30	au: eels	K
234656	48	kw: profumo, and kw: john	S
234656	38	su= "Profumo, John D."	X
234656	31	su= "Profumo, John D." and dt= "bks"	NT

234676	1	kw: chefs and kw: mukilteo	S
234676	0	kw: mukilteo'S and kw: ANSWER	MS
234677	16	(ti: Third and ti: Symphony) and au: Copland and dt= "rec"	K
234677	0	li: IVU and ((ti: Third and ti: Symphony) and au: Copland and dt= "rec")	K
234681	28	kw: teed and kw: off	S
234681	11	kw: teed and kw: off and dt= "bks"	NT
234690	4	kw: barbara and kw: longhi	S
234714	7	kw: tents and kw: architecture and kw: nomads	S
234716	0	au: john and au: seargant	K
234716	5	ti: usonian and ti: houses	K
234718	1	au: Toler, and au: Stanley	K
234718	1	ti: Seventy-five and ti: years and ti: powerful and ti: preaching	K
234792	4	ti: Organized and ti: Religion and ti: United and ti: States	K
234796	3	(ti: slow and ti: fade and ti: black) and ((au: cripps, and au: thomas))	K
234796	0	(ti: films and ti: oscar and ti: micheaux) and ((au: peterson, and au: bernard))	K
234796	66	au: peterson, and au: bernard	K
234796	9	au: peterson, and au: bernard and au: L.	K
234796	2	(ti: redefining and ti: black and ti: film) and ((au: reid, and au: mark and au: A.))	K
234796	3	((ti: blacks and ti: black) and ti: white) and au: sampson	K
234798	19	kw: joanne and kw: akalaitis	X
234805	0	ti: grey and ti: friars and ti: bobby	K
234805	0	ti: gray and ti: friars and ti: bobby	K
234805	0	ti: gray and ti: friars and ti: bobbie	K
234805	0	ti: grey and ti: friars and ti: bobbie	K
234805	76	kw: grey and kw: friars	S
234805	0	kw: grey and kw: friars and kw: bob	NS
234805	76	kw: grey and kw: friars	X
234805	71	kw: grey and kw: friars and dt= "bks"	NT
234805	56	kw: greyfriars and kw: bobby	X
234805	38	kw: greyfriars and kw: bobby and dt= "bks"	NC

234817	5	ti: OFFICIAL and ti: PREPPY and ti: HANDBOOK	K
234838	53	au: reger and ((kw: trio and kw: fur and kw: violine))	K
234838	1	no: 10434882	K
234838	2	ti: Trio and ti: für and ti: Violine, and ti: Viola and ti: und and ti: Klavier, and ti: opus and ti: 2	K
234842	2	kw: schaumburger and kw: auswanderer	S
234896	1715	kw: air and kw: traffic and kw: control and kw: systems	S
234896	66	su= "air traffic control systems"	X
234896	408	kw: air and kw: traffic and kw: control and kw: systems and yr: 1995-2002	NT
234896	245	su: air and su: traffic and su: control and su: systems and yr: 1995-2002	X
234910	6	ti: finding and ti: forrester	K
234912	10	au: alire, and au: camila	K
234912	0	ti: Diversity and ((ti: Leadershp and ti: Color and ti: Leadership))	K
234927	8	(kw: developing and kw: library) and ((kw: information and kw: center and kw: collections))	S
234934	316	ti: phylogenetic and ti: analysis	K
234934	13	ti: phylogenetic and ti: analysis and ti: morphological	K
234934	4	au= "Wiens, John J."	K
234938	75	au: schelle and au: michael	K
234938	10	au: wentzel and au: wayne	K
234939	0	kw: natowitz, and kw: allen	S
234939	0	au: natowitz, and au: allen	K
234948	15	kw: glam and kw: rock and (dt= "bks" or dt= "art")	S
234959	1	(ti: good and ti: fight) and ((au: brady, and au: sarah))	K
234959	0	(ti: why and ti: we and ti: fight) and ((au: bennett, and au: william))	K
234959	127	ti: why and ti: we and ti: fight	K
234959	49	ti: why and ti: we and ti: fight and dt= "bks"	K
234963	0	(ti: inorganic and ti: synthesis) and au: brauer and yr: -1985	K
234963	0	(ti: inorganic and ti: synthesis) and au: brouer and yr: -1985	K
234963	33	ti: inorganic and ti: synthesis and yr: -1985	K
234965	11	kw: ethiopia and kw: decentralization	S
234965	7	(su= "Decentralization in government") and su= "Ethiopia."	X

234965	344	au= "United Nations Centre for Regional Development."	K
234965	15	su= "united nations centre for regional development"	S
234968	22	au: DEMIMUID, and au: MAURICE	K
234968	106	au: SLEATOR, and au: WILLIAM	K
234976	0	nb: 1558290737	K
234976	2	ti: dake and ti: annotated and ti: reference and ti: bible	K
234976	1	ti= "dake annotated reference bible"	K
234980	6	kw: grimm and ((kw: lettres and kw: inedites))	S
234980	4	ti: grimm and ((ti: lettres and ti: inedites))	K
234984	5	kw: amor y kw: pecado and dt= "vis"	S
234984	2	kw: el and kw: guardian and kw: de and kw: la and kw: muerte and dt= "vis"	MS
234984	1	kw: un and kw: luz and kw: en and kw: la and kw: escalera and dt= "vis"	MS
234984	3	kw: la and kw: negra and kw: angustias and dt= "vis"	MS
234998	0	ti: Ramon and ti: empieza and ti: el and ti: curso and dt= "ser"	K
234998	0	ti: Ramon and ti: comienza and ti: el and ti: curso and dt= "ser"	K
234998	0	ti: Ramons and ti: starts and ti: school and dt= "ser"	K
235010	1	ti: gee and ti: family and ti: tree	K
235010	1	ti: story and ti: louis and ti: tetreau	K
235010	73	ti: wayne and ti: county and ti: west and ti: virginia	K
235018	342	kw: "the" and kw: elliotts	S
235018	0	(kw: "the" and kw: elliotts) and kw: bbc	NT
235018	3	(kw: elliott and kw: house) and kw: bbc	MC
235025	0	(ti: India's and ti: foreign and ti: policy) and kw: vikas	K
235025	156	ti: India's and ti: foreign and ti: policy	K
235025	11	ti: India's and ti: foreign and ti: policy and yr: 1984-1985 and dt= "bks"	K
235040	1	ti: nothing and ti: makes and ti: free	K
235041	4	au: bucher, and au: ward	K
235041	6	au: gowans, and au: ann	K
235041	46	au: gowans, and au: alan	K
235041	4	au: hutt, and au: sherry	K

235041	28	au: lee, and au: antoinette	K
235041	4	au: hutt, and au: sherry	K
235041	0	au: moore, and au: authur and au: c	K
235041	0	au: moore, and au: authur	K
235041	883	au: moore, and au: arthur	K
235041	13	(au: moore, and au: arthur) and kw: preservation	K
235041	6	au: smeallie, and au: peter	K
235041	38	ti: architecture and (ti: the and ti: city) and dt= "vis"	K
235041	2	ti: architecture and (ti: the and ti: city) and yr: 1988 and dt= "vis"	K
235041	7	au: barthel, and au: diane	K
235041	6	au: driggs, and au: sarah	K
235041	1	(au: graham, and au: brian) and kw: preservation	K
235041	1	(au: hamer, and au: david) and kw: preservation	K
235041	1	(au: graham, and au: brian) and kw: preservation	K
235041	51	au: hardesty, and au: donald	K
235041	2	(au: kelley, and au: stephen) and kw: preservation	K
235041	3	(au= "Kelley, Stephen J.,") and au= "1954-"	K
235041	2	(au: strike, and au: james) and kw: preservation	K
235041	17	au: strike, and au: james	K
235041	4	au: weyeneth, and au: robert	K
235041	4	au: bucher, and au: ward	K
235041	46	au: gowans, and au: alan	K
235041	2	au: gowans, and au: alan and dt= "vis"	K
235054	1	nb: 1779010486	K
235054	1	nb: 996684646	K
235054	1	nb: 997002146x	K
235054	1	nb: 9964720106	K
235054	1	nb: 9781299541	K
235054	1	nb: 9970021532	K
235054	2	nb: 9780290141	K

235054	1	nb: 0908311826	K
235058	3	(ti: island and ti: sun) and ((au: villoldo, and au: alberto))	K
235072	682	kw: nudes	S
235072	481	kw: nudes and dt= "bks"	NT
235100	65	au: griggs, and au: donald and au: I	K
235113	35	kw: ben and kw: hogan and dt= "vis"	S
235118	16	kw: defects and kw: silicon and kw: III	S
235118	1	(au= "International Symposium on Defects in Silicon") AND au= "1999" and (au= "Seattle, Wash.")	K
235121	0	nb: 1884590772	K
235121	1	nb: 1884590780	K
235140	141	kw: "the" and kw: impact and kw: historic and kw: preservation	S
235140	4	kw: "the" and kw: impact and kw: historic and kw: preservation and li: IWA	X
235140	375	kw: rural and kw: industry, and kw: economic and kw: development	S
235140	8	kw: rural and kw: industry and kw: revitalization	MS
235140	30	(su= "Rural renewal") and (su= "United States.")	MS
235140	381	(su= "Small business") AND (su= "United States") and (su= "Handbooks, manuals, etc.")	MS
235140	11	au= "Borich, Timothy O."	K
235140	0	kw: rural and kw: renewel	MS
235140	0	kw: rural and kw: renewl	X
235140	0	kw: rural and kw: renewl and kw: development	MS
235140	106	kw: rural and kw: renewal and kw: planning	MS
235140	7	kw: rural and kw: renewal and kw: planning and li: IWA	X
235140	299	su= "rural renewal"	X
235140	0	kw: rural and kw: industry and kw: development,renewal and li: IWA	MS
235140	0	kw: rural and kw: industry and kw: revitalizaiton and li: IWA	X
235140	178	kw: rural and kw: industry and li: IWA	BS
235143	3	kw: academic and kw: psychiatry and dt= "ser"	S
235143	12	(su= "Psychiatry" AND su= "education") and su= "periodicals."	X

APPENDIX C

Distributions of Data Sets

(i) For Raw Index

World	NonFt	Fulltext	Group B	Group Z	OCLCArt	ABI	BusMan	EdAbs	HWI	Med	PerAbs	SocAbs
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	4	0	0	0	0	0	0	0	0	0
0	0	0	5	0	0	0	0	0	0	0	0	0
0	0	0	7	0	0	0	0	0	0	0	0	0
0	0	0	14	0	0	0	0	0	0	0	0	1
0	0	0	20	0	0	1	0	0	0	0	0	1
0	0	0		0	0	1	0	0	0	0	0	1
0	0	0		0	0	1	0	0	0	0	0	1
0	0	0		0	1	2	1	1	0	0	0	1
0	0	0		0	2	2	1	1	0	0	0	1
0	0	0		0	3	2	1	1	0	0	0	1
0	0	0		0	4	2	1	1	0	0	1	1
0	0	0		0	5	3	1	1	0	0	1	2
0	0	0		0	8	4	1	1	0	0	1	2
0	0	0		0	15	5	1	2	0	0	1	3

0	0	0	0	28	7	1	2	0	0	1	3
0	0	0	0		9	1	2	0	0	2	3
0	0	0	0		13	1	2	0	0	2	3
0	0	0	0		13	1	2	0	0	2	4
0	0	0	0		15	2	2	0	0	2	5
0	0	0	0			2	2	0	0	2	5
0	0	0	1			2	2	0	0	3	7
0	0	0	1			2	3	0	0	3	9
0	0	0	1			2	4	0	0	3	10
0	0	0	1			2	4	0	0	3	16
0	0	0	1			2	5	0	0	4	35
0	0	0	1			2	5	0	0	5	
0	0	0	1			2	7	0	0	5	
0	0	0	1			3	12	0	0	6	
0	0	0	1			3	21	0	0	9	
0	0	0	1			3		0	0	9	
0	0	0	1			3		0	0	9	
0	0	0	1			3		0	0	15	
0	0	0	2			4		0	0	21	
0	0	0	2			4		0	0		
0	0	0	2			4		0	0		
0	0	0	2			4		0	0		
0	0	0	2			5		0	0		
0	0	0	3			5		0	0		
0	0	0	3			5		0	0		
0	0	0	3			6		0	0		
0	0	0	3			6		0	0		
0	0	0	3			7		0	0		
0	0	0	3			7		0	0		
0	0	0	3			7		0	0		
0	0	0	3			8		0	0		

0	0	0	4	8	0	0
0	0	0	4	9	0	1
0	0	0	4		0	1
0	0	0	5		0	1
0	0	0	5		0	1
0	0	0	6		0	1
0	0	0	7		0	1
0	0	0	8		0	1
0	0	0	8		0	1
0	0	0	10		0	1
0	0	0	12		0	1
0	0	0	14		0	1
0	0	0	20		0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			0	1
0	0	0			1	1
0	0	0			1	1
0	0	0			1	1
0	0	0			1	1
0	0	0			1	1

0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2

0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	2	2
0	0	0	2	2
0	0	0	2	3
0	0	0	2	3
0	0	0	2	3
0	0	0	2	3
0	0	1	2	3
0	0	1	2	3
0	0	1	2	3
0	0	1	2	3
0	0	1	2	3
0	0	1	2	3
0	0	1	2	3
0	0	1	2	3
0	0	1	2	3
0	0	1	2	4
0	0	1	2	4
0	0	1	2	4
0	0	1	2	4
0	0	1	2	4
0	0	1	2	5
0	0	1	2	5
0	0	1	2	5
0	0	1	2	5
0	0	1	2	6
0	0	1	2	6
0	0	1	2	6
0	0	1	2	6

0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
2
2
2
2
2
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
4
4
4
4
4
4
4
4
4
4
4
4
5
5
5
5
6
6
7
7
7
8
8
10
11
11
11
12
13
17
18
23

0	0	2	5
0	0	2	5
0	0	2	5
0	0	2	5
0	0	2	5
0	0	2	5
0	0	2	6
0	0	2	6
0	0	2	6
0	0	2	6
0	0	2	6
0	0	2	6
0	0	2	6
0	0	2	7
0	0	2	7
0	0	2	7
0	0	2	7
0	0	2	7
0	0	2	7
0	0	2	7
0	0	2	7
0	0	2	8
0	0	2	8
0	0	2	8
0	0	2	8
0	0	2	8
0	0	2	9
0	0	2	9
0	0	3	9
0	0	3	10
0	0	3	10

0	0	3	10
0	0	3	10
0	0	3	12
0	0	3	12
0	0	3	13
0	0	3	13
0	0	3	13
0	0	3	14
0	0	3	16
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	4	
0	0	4	
0	0	4	
0	0	4	
0	0	4	

0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	5
0	0	5
0	0	5
1	0	5
1	0	5
1	0	5
1	0	5
1	0	5
1	0	5
1	0	5
1	0	6
1	0	6
1	0	6
1	0	6
1	0	6
1	0	6
1	0	6
1	0	6
1	0	6
1	0	6
1	0	6
1	0	7
1	0	7
1	0	7
1	0	8

1	0	8
1	0	8
1	0	9
1	0	9
1	0	9
1	0	9
1	0	9
1	0	9
1	0	10
1	0	10
1	0	12
1	0	12
1	0	14
1	0	16
1	0	29
1	0	30
1	0	33
1	0	34
1	0	
1	0	
1	0	
1	0	
1	0	
1	0	
1	0	
1	0	
1	0	
1	0	
1	0	
1	0	
1	0	

1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
2	0
2	0

2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1

2	1
2	1
2	1
2	1
2	1
2	1
2	1
2	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
4	1
4	1

4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
5	1
5	1
5	1
5	1
5	1
5	1
5	1
5	1
5	1
6	1
6	1
6	1
6	1
6	1
6	1

7	1
8	1
8	1
8	1
11	1
15	1
15	1
15	1
19	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1

1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
2
2
2
2
2
2
2
2
2
2
2

2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2

2
2
2
2
2
2
2
2
2
2
2
2
2
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3

3
3
3
3
3
3
3
3
3
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4

4
5
5
5
5
5
5
5
5
5
5
5
5
5
5
6
6
6
6
6
6
6
6
7
7
7
7
7
7
8
8
8

8
9
9
9
9
9
9
9
10
11
12
14
14
16
17
18
27

(ii) For Refined Index

World	NonFt	Fulltext	Group B	Group Z	OCLCArt	ABI	BusMan	EdAbs	HWI	Med	PerAbs	SocAbs
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0

0	0	0	2	0	0	0	0	0	0	0	0	0
0	0	0	3	0	0	0	0	0	0	0	0	0
0	0	0	3	0	0	0	0	0	0	0	0	0
0	0	0	6	0	0	0	0	0	0	0	0	0
0	0	0	19	0	0	0	0	0	0	0	0	0
0	0	0		0	0	1	0	0	0	0	0	0
0	0	0		0	1	1	0	0	0	0	0	1
0	0	0		0	3	1	1	1	0	0	0	2
0	0	0		0	4	2	1	1	0	0	0	2
0	0	0		0	4	2	1	1	0	0	0	2
0	0	0		0	8	2	1	1	0	0	0	3
0	0	0		0	14	3	1	2	0	0	0	3
0	0	0		0	26	3	1	2	0	0	1	3
0	0	0		0		5	2	2	0	0	1	3
0	0	0		0		7	2	2	0	0	1	4
0	0	0		0		9	2	2	0	0	1	4
0	0	0		0		10	2	3	0	0	1	4
0	0	0		0		13	2	4	0	0	1	7
0	0	0		0		15	2	4	0	0	2	10
0	0	0		0			2	4	0	0	2	32
0	0	0		0			2	21	0	0	2	
0	0	0		0			2		0	0	3	
0	0	0		0			2		0	0	3	
0	0	0		0			3		0	0	3	
0	0	0		0			3		0	0	4	
0	0	0		0			3		0	0	5	
0	0	0		0			3		0	0	5	
0	0	0		1			3		0	0	9	

0	0	0	1	4	0	0	9
0	0	0	1	4	0	0	13
0	0	0	1	5	0	0	21
0	0	0	1	5	0	0	
0	0	0	1	6	0	0	
0	0	0	1	6	0	0	
0	0	0	2	7	0	0	
0	0	0	2	7	0	0	
0	0	0	2	8	0	0	
0	0	0	2	9	0	0	
0	0	0	2		0	0	
0	0	0	2		0	0	
0	0	0	2		0	0	
0	0	0	2		0	0	
0	0	0	3		0	0	
0	0	0	3		0	0	
0	0	0	3		0	0	
0	0	0	3		0	0	
0	0	0	3		0	0	
0	0	0	3		0	0	
0	0	0	5		0	0	
0	0	0	5		0	0	
0	0	0	6		0	0	
0	0	0	7		0	0	
0	0	0	8		0	0	
0	0	0	8		0	0	
0	0	0	9		0	0	
0	0	0	11		0	0	

0	0	0	14	0	0
0	0	0	19	0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		0	1
0	0	0		1	1
0	0	0		1	1
0	0	0		1	1
0	0	0		1	1
0	0	0		1	1

0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	2
0	0	0	1	3
0	0	0	2	3
0	0	0	2	3
0	0	0	2	3
0	0	0	2	3
0	0	0	2	3
0	0	0	2	3

0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	1
0	0	1
0	0	1
0	0	1
0	0	1
0	0	1
0	0	1
0	0	1
0	0	1
0	0	1
0	0	1
0	0	1
0	0	1
0	0	1

2	4
2	4
2	4
2	4
2	4
2	5
2	5
2	5
2	6
2	6
2	6
2	6
2	7
2	8
2	9
3	11
3	11
3	11
3	12
3	13
3	16
3	16
3	22
3	
3	
4	
4	
4	

0	0	1	4
0	0	1	4
0	0	1	4
0	0	1	4
0	0	1	4
0	0	1	4
0	0	1	4
0	0	1	4
0	0	1	4
0	0	1	5
0	0	1	5
0	0	1	5
0	0	1	5
0	0	1	5
0	0	1	5
0	0	1	5
0	0	1	5
0	0	1	5
0	0	1	5
0	0	1	6
0	0	1	6
0	0	1	6
0	0	1	6
0	0	1	6
0	0	1	6
0	0	1	6
0	0	1	6
0	0	1	7
0	0	1	7
0	0	1	7

0	0	1	7
0	0	1	7
0	0	1	8
0	0	1	8
0	0	1	8
0	0	1	8
0	0	1	9
0	0	1	9
0	0	1	10
0	0	2	10
0	0	2	11
0	0	2	12
0	0	2	13
0	0	2	14
0	0	2	16
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	

0	0	2
0	0	2
0	0	2
0	0	2
0	0	2
0	0	2
0	0	2
0	0	2
0	0	2
0	0	2
0	0	2
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3
0	0	3

0	0	3
0	0	3
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	4
0	0	5
0	0	5
0	0	5
0	0	5
0	0	5
0	0	5
0	0	5
0	0	5
0	0	5
0	0	6
0	0	6

0	0	6
0	0	6
0	0	6
0	0	6
0	0	6
0	0	6
0	0	7
0	0	7
0	0	7
0	0	7
0	0	7
0	0	8
0	0	8
0	0	8
0	0	8
0	0	9
0	0	9
0	0	9
0	0	10
0	0	10
0	0	12
0	0	12
0	0	14
0	0	16
0	0	27
0	0	30
0	0	33
0	0	33

0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0
0	0

0	0
0	0
0	0
0	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0

1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
1	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0

2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
2	0
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1
3	1

3	1
3	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
4	1
5	1
5	1
5	1
5	1
5	1
5	1
5	1
5	1
5	1
6	1
6	1
7	1
8	1

8	1
9	1
12	1
14	1
15	1
17	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1

1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1

1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
1
2
2
2
2
2
2
2
2
2
2
2
2

2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2
2

2
2
2
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3
3

3
3
3
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
4
5
5
5
5
5
5
5
5
5
5
5

5
5
6
6
6
6
6
6
6
6
6
6
7
7
7
8
8
8
9
9
9
9
9
10
11
11
13
14

15
16
17
27

(iii) Control Data Sets

(a) Control distributions by raw index

World30	World31	Med28	Med29
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	1	1
0	0	1	1
0	0	1	1
0	0	1	2
0	0	1	2
0	0	1	3

0	0	1	3
0	0	1	3
0	0	1	3
0	0	1	4
0	0	1	4
0	0	1	7
0	0	1	8
0	0	1	10
0	0	1	11
0	0	1	18
0	0	1	
0	0	1	
0	0	1	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	3	
0	0	5	
0	0	5	
0	0	5	
0	0	6	
0	0	6	
0	0	7	

0	0	8
0	0	11
0	0	13
0	0	16
0	0	23
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	1	
0	1	
0	1	
0	1	
0	1	
0	1	
0	1	
0	1	
0	1	
0	1	
0	1	
1	1	
1	1	

1	1
1	1
1	1
1	1
1	1
1	1
1	1
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
2	3
2	3
2	3
2	3

2	3
2	4
2	4
2	4
2	4
2	5
2	6
2	7
2	10
3	10
3	11
3	12
3	
3	
3	
3	
3	
3	
3	
3	
4	
4	
4	
4	
5	
5	
6	
6	

7
8
8
9

(b) Control distributions by refined index

World30	World31	Med28	Med29
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	1
0	0	0	2
0	0	0	2
0	0	0	3

0	0	0	3
0	0	0	3
0	0	0	3
0	0	0	4
0	0	1	4
0	0	1	4
0	0	1	5
0	0	1	9
0	0	1	11
0	0	1	12
0	0	1	
0	0	1	
0	0	1	
0	0	1	
0	0	1	
0	0	1	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	2	
0	0	3	
0	0	5	
0	0	6	
0	0	6	
0	0	6	
0	0	7	
0	0	11	
0	0	13	

0	0	16
0	0	22
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	
0	0	

0	0
0	0
0	0
0	0
0	0
0	0
0	1
0	1
0	1
0	1
0	1
0	1
0	1
0	1
0	1
0	1
1	1
1	1
1	1
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	2
1	3
1	4

1	4
2	7
2	9
2	
2	
2	
2	
2	
2	
2	
2	
2	
3	
3	
3	
3	
3	
3	
3	
3	
6	
7	
7	

REFERENCES

Aigrain, P. & Longueville, V. (1994) "A model for the evaluation of expansion techniques in information-retrieval systems" *Journal of the American Society for Information Science* 45 (4):225-234.

Allen, B.L. (1993) "Improved browsable displays: an experimental test" *Information Technology and Libraries* 12 (2):203-208.

Ballard, T. (1994) "Comparative searching styles of patrons and staff" *Library Resources & Technical Services* 38:293-305.

Ballard, T. & Smith, J. (1992) "The human interface: an ongoing study of OPAC usage at Adelphi University" in *Advances in Online Public Access Catalogs. 1,* edited by M. Ra, pp. 58-73 (Westport: Meckler).

Barry, C.L. (1994) "User-defined relevance criteria: an exploratory study" *Journal of the American Society for Information Science* 45 (4):149-159.

Barry, C.L. (1998) "Documentation representations and clues to document relevance" *Journal of the American Society for Information Science* 49 (14):1293-1303.

Barry, C.L. & Schamber, L. (1998) "Users' criteria for relevance evaluation: a cross-situational comparison" *Information Processing & Management* 34 (2-3):219-236.

Bates, M.J. (1989) "The design of browsing and berrypicking techniques for the online search interface" *Online Review* 13 (5):407-424.

Bawden, D. (1986) "Information systems and the stimulation of creativity" *Journal of Information Science* 12 (5):203-216.

Bawden, D. (1993) "Browsing: theory and practice" *Perspectives in Information Management* 3 (1):71-85.

Beaulieu, M. (2000) "Interaction in information searching and retrieval" *Journal of Documentation* 56 (4):431-439.

Beheshti, J. (1997) "The evolving OPAC" *Cataloging & Classification Quarterly* 24 (1/2):163-185.

Beheshti, J., Large, V. & Bialek, M. (1996) "PACE: a browsable graphical interface" *Information Technology and Libraries* 15 (4):231-240.

Belkin, N.J. (1980) "Anomalous states of knowledge as a basis for information retrieval" *Canadian Journal of Information Science* 5:33-143.

Belkin, N.J., Marchetti, P.G. & Cool, C. (1993) "Braque: design of an interface to support user interaction in information retrieval" *Information Processing & Management* 29 (3):325-344.

Blecic, D.D., Bangalore, N.S. & Dorsch, J.L. (1998) "Using transaction log analysis to improve OPAC retrieval results" *College & Research Libraries* 59 (1):39-50.

Blecic, D.D. and others (1999) "A longitudinal study of the effects of OPAC screen changes on searching behavior and searcher success" *College & Research Libraries* 60 (6):515-530.

Borgman, C.L. (1996) "Why are online catalogs still hard to use?" *Journal of the American Society for Information Science* 47 (7):493-503.

Brookes, B.C. (1981) "Information technology and the science of information" in *Information retrieval research,* edited by R.N. Oddy and others, pp. 1-8 (London: Butterworths).

Bruce, H.W. (1994) "A cognitive view of the situational dynamism of user-centered relevance estimation" *Journal of the American Society for Information Science* 45 (3):142-148.

Chang, S-J. & Rice, R.E. (1993) "Browsing: a multidimensional framework" *Annual Review of Information Science & Technology* 28:231-276.

Cherry, J.M. (1992) "Improving subject access in OPACs: an exploratory study of conversion of users' queries" *Journal of Academic Librarianship* 18:95-99.

Connaway, L.S., Budd, J.M. & Kochtanek, T.R. (1995) "An investigation of the use of an online catalog: user characteristics and transaction log analysis" *Library Resources & Technical Services* 39:142-152.

Cooper, William S. (1973) "On selecting a measure of retrieval effectiveness" *Journal of the American Society for Information Science* 24 (2):87-100.

Cremmins, E.T. (1992) "Value-added processing of representational and speculative information using cognitive skills" *Journal of Information Science* 18 (1):27-43.

Croft, W.B. & Thompson, R.H. (1987) "I^3R: a new approach to the design of document retrieval systems" *Journal of the American Society for Information Science* 38 (6):389-404.

Cronin, B. & Hert, C.A. (1995) "Scholarly foraging and network discovery tools" *Journal of Documentation* 51:388-403.

Dalrymple, P.W. (1990) "Retrieval by reformulation in two library catalogs: toward a cognitive model of searching behavior" *Journal of the American Society for Information Science* 41:272-281.

Debili, F., Fluhr, C. & Radasoa, P. (1989) "About reformulation in full-text IRS" *Information Processing & Management* 25 (6):647-657.

Dervin, B. (1977) "Useful theory for librarianship: communication, not information" *Drexel Library Quarterly* 13 (3):16-32.

Doll, R. (1980) "Speculations on unanticipated outcomes of data base research" *Journal of Library Automation* 13 (1):49-55.

Eisenberg, M. & Barry, C. (1988) "Order effects: a study of the possible influence of presentation order on user judgments of document relevance" *Journal of the American Society for Information Science* 39 (5):293-300.

Ellis, D. (1989) "A behavioral approach to information retrieval system design" *Journal of Documentation* 45 (3):171-212.

Erdelez, S. (1997) "Information encountering: a conceptual framework for accidental information discovery" in *Information seeking in context: proceedings of an International Conference on Research in Information Needs, Seeking and Use in Different Contexts, 14-16 August, 1996, Tampere, Finland,* edited by P. Vakkari, R. Savolainen & B. Dervin, pp. 412-421 (London: Taylor Graham).

Flaherty, P. (1993) "Transaction logging systems: a descriptive summary" *Library Hi Tech* 11 (2):67-78.

Frarey, C.J. (1953) "Studies of the use of the subject catalog: summary and evaluation" in *The subject analysis of library materials,* edited by M.F. Tauber, pp. 147-166 (New York: Columbia University).

Greisdorf, H. & Spink, A. (2000) "Recent relevance research: implications for information professionals" *Online Information Review* 24 (5):389-395.

Greisdorf, H. & Spink, A. (2001) "Median measure: an approach to IR systems evaluation" *Information Processing & Management* 37 (6):843-857.

Hafter, R. (1979) "Performance of card catalogs: a review of research" *Library Research* 1 (3):199-222.

Haglund, H. (1996) "The three princes of Serendip and the cyberspace" *FID News Bulletin* 46 (7/8):240-243.

Hancock-Beaulieu, M. (1990) "Evaluating the impact of an online library catalogue on subject searching behaviour at the catalogue and at the shelves" *Journal of Documentation* 46 (4):318-338.

Hancock-Beaulieu, M. (1993) "A comparative transaction log analysis of browsing and search formulation in online catalogues" *Program* 27:269-280.

Hancock-Beaulieu, M., Fieldhouse, M. & Do, T. (1995) "An evaluation of interactive query expansion in an online library catalogue with a graphical user interface" *Journal of Documentation* 51 (3):225-243.

Harmon, G. (1970) "Information science as an integrative discipline" *Proceedings of the American Society for Information Science* 6:40-42.

Harmon, E.G. & Ballesteros, E.R. (1997) "Unconscious cognition: the elicitation of deeply embedded information needs" in *Information seeking in context: proceedings of an International Conference on Research in Information Needs, Seeking and Use in Different Contexts, 14-16 August, 1996, Tampere, Finland,* edited by P. Vakkari, R. Savolainen & B. Dervin, pp. 422-433 (London: Taylor Graham).

Harter, S.P. (1992) "Psychological relevance and information science" *Journal of the American Society for Information Science* 43 (9):602-615.

Herner, S. (1970) "Browsing" in *Encyclopedia of library and information science* 3:408-415.

Hert, C.A. (1995) "Information retrieval as situated action" in *ASIS '95: proceedings of the 58th ASIS annual meeting,* edited by T. Kinney, pp. 172-180 (Medford: Information Today).

Hert, C.A. (1997) *Understanding information retrieval interactions: theoretical and practical implications* (Greenwich: Ablex).

Hildreth, C.R. (1997) "The use and understanding of keyword searching in a university online catalog" *Information Technology & Libraries* 16:52-62.

Hillman, D.J. (1968) "Negotiation of inquiries in an on-line retrieval system" *Information Storage & Retrieval* 4 (2):219-238.

Hjørland, B. (1997) *Information seeking and subject representation: an activity-theoretical approach to information science* (Westport: Greenwood Press).

IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998) *Functional requirements for bibliographic records* (Munchen: K.G. Saur). <http://www.ifla.org/VII/s13/frbr/frbr.pdf>

Ingwersen, P. (1992) *Information retrieval interaction* (London: Taylor Graham).

Ingwersen, P. (1996) "Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory" *Journal of Documentation* 52 (1):3-50.

Janes, J.W. (1991) "Relevance judgments and the incremental presentation of document representations" *Information Processing & Management* 27 (6):629-646.

Jansen, B.J. & Spink, A. (2000) "Methodological approach in discovering user search patterns through Web log analysis" *Bulletin of the American Society for Information Science* 27 (1):15-17.

Jansen, B.J., Spink, A. & Saracevic, T. (2000) "Real life, real users, and real needs: a study and analysis of user queries on the web" *Information Processing & Management* 36:207-227.

Järvelin, K. & Wilson, T.D. (2003) "On conceptual models for information seeking and retrieval research" *Information Research* 9 (1), paper 163 <http://InformationR.net/ir/9-1/paper163.html>

Jones, S., Gatford, M. & Do, T. (1997) "Transaction logging in Okapi-related projects" *Journal of Documentation* 53 (1):35-50.

Kanji, G.K. (1999) *100 statistical tests,* 2nd ed. (London: Sage).

Katzer, J. & Snyder, H. (1990) "Toward a more realistic assessment of information retrieval performance" in *ASIS '90: proceedings of the 53rd ASIS Annual Meeting,* edited by Diane Henderson, pp. 80-85 (Medford: Information Today).

Korfhage, R. (1991) "To see, or not to see: is that the query?" in *SIGIR '91: proceedings of the Fourteenth Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval,* edited by A. Bookstein and others, pp. 134-141 (New York: ACM).

Kranich, N.C. and others (1986) "Evaluating the online catalog from a public services perspective: a case study at the New York University Libraries" in *The impact of online catalogs,* edited by J.R. Matthews, pp. 89-140 (New York: Neal-Schuman).

Kuhlthau, C.C. (1991) "Inside the search process: information seeking from the user's perspective" *Journal of the American Society for Information Science* 42 (5):361-371.

Kuhlthau, C. & Cool, C. (1992) "Exploration into stages in the information search process in online information retrieval: communication between users and intermediaries" in *ASIS '92: proceedings of the 55th ASIS annual meeting,* edited by D. Shaw, pp. 67-71 (Medford: Information Today).

Large, A. & Beheshti, J. (1997) "OPACs: a research review" *Library & Information Science Research* 19:111-133.

Larson, R.R. (1991) "The decline of subject searching: long-term trends and patterns of index use in an online catalog" *Journal of the American Society for Information Science* 42 (3):197-215.

Liestman, D. (1992) "Chance in the midst of design: approaches to library research serendipity" *RQ* 31 (4):524-532.

Lin, X. (1995) "Searching and browsing on map displays" in *ASIS '95: proceedings of the 58th ASIS annual meeting,* edited by T. Kinney, pp. 13-18 (Medford: Information Today).

Lin, X. (1997) "Map displays for information retrieval" *Journal of the American Society for Information Science* 48 (1):40-54.

Lipetz, B.-A. (1970) *User requirements in identifying desired works in a large library* (Washington, D.C.: U.S. Office of Education, Bureau of Research).

Marcus, R.S., Kugel, P. & Benenfeld, A.R. (1978) "Catalog information and text as indicators of relevance" *Journal of the American Society for Information Science* 29 (1):15-30.

Martin, T.H., Wyman, J.C. & Madlock, K. (1983) *Feedback and exploratory mechanisms for assisting library staff improve on-line catalog searching* (Washington, D.C.: Council on Library Resources).

Matthews, J.R. (ed.) (1986) *The impact of online catalogs* (New York: Neal-Schuman).

Mizzaro, S. (1998) "How many relevances in information retrieval?" *Interacting with Computers* 10 (3):303-320.

Ng, K.B. (2002) "Toward a theoretical framework for understanding the relationship between situated action and planned action models of behavior in information retrieval contexts: contributions from phenomenology" *Information Processing & Management* 38 (5):613-626.

Nicholas, D. (1996) "An assessment of the online searching behaviour of practitioner end users" *Journal of Documentation* 52 (3):227-251.

O'Brien, A. (1994) "Online catalogs" *Annual Review of Information Science & Technology* 29:219-242.

Olsen, K.A., Korfhage, R.R. & Sochats, K.M. (1993) "Visualization of a document collection: the VIBE system" *Information Processing & Management* 29 (1):69-81.

Ozmutlu, H.C., Spink, A. & Ozmutlu, S. (2002) "Analysis of large data logs: an application of Poisson sampling on Excite Web queries" *Information Processing & Management* 38 (4):473-490.

Paley, A.J. & Fox, M.S. (1981) "Browsing through databases" in *Information retrieval research,* edited by R.N. Oddy and others, pp. 310-324 (London: Butterworths).

Pao, M.L. (1989) *Concepts of information retrieval* (Englewood: Libraries Unlimited).

Pejtersen, A.M. (1993) "A new approach to design of document retrieval and indexing systems for OPAC users" in *Online Information 93,* edited by D. Raitt & B. Jeapes, pp. 273-290 (Oxford: Learned Information).

Peters, T.A. (1989) "When smart people fail: an analysis of the transaction log of an online public access catalog" *Journal of Academic Librarianship* 15:267-273.

Peters, T.A. (1993) "The history and development of transaction log analysis" *Library Hi Tech* 110:41-66.

Pirolli, P. & S. Card. (1999) "Information foraging" *Psychological Review* 106 (4):643-675.

Powell, R.R. (1999) "Recent trends in research: a methodological essay" *Library & Information Science Research* 21 (1):91-119.

Pu, H.-T., Chuang, S.-L. & Yang, C. (2002) "Subject categorization of query terms for exploring Web users' search interests using Taiwan search engine logs" *Journal of the American Society for Information Science & Technology* 53 (8):617-630.

Ray, K.L. & Lang, M.S (1997) "Analyzing search styles of patrons and staff: a replicative study of two university libraries" *Library Resources & Technical Services* 41:219-235.

Rice, J. (1988) "Serendipity and holism: the evolving OPAC" *Library Journal* 113 (3):138-141.

Richardson, R.J. (1981) "End-user online searching in a high-technology engineering environment" *Online* 5 (4):44-57.

Robertson, S.E. & Hancock-Beaulieu, M.M. (1992) "On the evaluation of IR systems" *Information Processing & Management* 28 (4):457-466.

Robins, D. (1997) "Shifts of focus in information retrieval interaction" in *ASIS '97: proceedings of the 60th ASIS Annual Meeting,* edited by C. Schwartz & M. Rorvig, pp. 123-134 (Medford: Information Today).

Robins, D. (2000) "Shifts of focus on various aspects of user information problems during interactive information retrieval" *Journal of the American Society for Information Science* 51 (10):913-928.

Sandstrom, P.E. (2001) "Scholarly communication as a socioecological system" *Scientometrics* 51 (3):573-605.

Saracevic, T. (1969) "Comparative effects of titles, abstracts and full texts on relevance judgements" *Proceedings of the American Society for Information Science* 6:293-299.

Saracevic, T. (1996) "Modeling interaction in information retrieval (IR): a review and proposal" in *ASIS '96: proceedings of the 59th ASIS Annual Meeting,* edited by S. Hardin, pp. 3-9 (Medford: Information Today).

Saracevic, T. (1997) "The stratified model of information retrieval interaction: extension and applications" in *ASIS '97: proceedings of the 60th ASIS Annual Meeting,* edited by C. Schwartz & M. Rorvig, pp. 313-327 (Medford: Information Today).

Schamber, L. (1994) "Relevance and information behavior" *Annual Review of Information Science & Technology* 29:3-48.

Shenouda, W. (1990) "Online bibliographic searching: how end-users modify their search strategies" in *ASIS '90: proceedings of the 53rd ASIS Annual Meeting,* edited by Diane Henderson, pp. 117-128 (Medford: Information Today).

Sloan, S.K. (1991) "Research and information retrieval among academic researchers: implications for library instruction" *Library Trends* 39:238-258.

Smithson, S. (1994) "Information retrieval evaluation in practice: a case study approach" *Information Processing & Management* 30:205-221.

Snelson, P. (1994) "Remote users of OPACs: do they differ from library users?" in *IOLS '94: proceedings of the Ninth Annual Conference on Integrated Online Library Systems*, edited by P.R. Cibbarelli, pp. 179-183 (Medford: Learned Information).

Spink, A. (1997) "Study of interactive feedback during mediated information retrieval" *Journal of the American Society for Information Science* 48 (5):225-250.

Spink, A. & Beatty, L. (1995) "Multiple search sessions by end-users of online catalogs and CD-ROM databases" in *ASIS '95: proceedings of the 58th ASIS annual meeting,* edited by T. Kinney, pp. 77-83 (Medford: Information Today).

Spink, A. & Greisdorf, H. (1997) "Partial relevance judgments and changes in users' information problems during online searching" in *Proceedings of the 18th National Online Meeting,* edited by M.E. Williams, pp. 323-334 (Medford: Information Today).

Spink, A. & Greisdorf, H. (2001) "Regions and levels: measuring and mapping users' relevance judgments" *Journal of the American Society for Information Science & Technology* 52 (2):161-173.

Spink, A., Greisdorf, H. & Bateman, J. (1998) "From highly relevant to not relevant: examining different regions of relevance" *Information Processing & Management* 34 (5):599-621.

Spink, A., Jansen, B.J. & Ozmutlu, C.H. (2000) "Use of query reformulation and relevance feedback by Excite users" *Internet Research* 10 (4):317-328.

Su, L.T. (1998) "Value of search results as a whole as the best single measure of information retrieval performance" *Information Processing & Management* 34 (5):557-579.

Sutcliffe, A.G. & Ennis, M. (1998) "Towards a cognitive theory of information retrieval" *Interacting with Computers* 10 (3):321-351.

Sutton, S.A. (2001) "The role of attorney mental models of law in case relevance determinations: an exploratory analysis" *Journal of the American Society for Information Science & Technology* 43 (3):186-200.

Tague-Sutcliffe, J. (1992) "Measuring the informativeness of a retrieval process" in *SIGIR '92: proceedings of the Fifteenth Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval,* edited by N. Belkin, P. Ingwersen & A.M. Pejtersen, pp. 23-36 (New York: ACM).

Tang, R., Shaw, W.M. & Vevea, J.L. (1999) "Towards the identification of the optimal number of relevance categories" *Journal of the American Society for Information Science* 50 (3):254-264.

Tang, R. & Solomon, P. (1998) "Toward an understanding of the dynamics of relevance judgment: an analysis of one person's search behavior" *Information Processing & Management* 34 (2/3):237-256.

Tang, R. & Solomon, P. (2001) "Use of relevance criteria across stages of document evaluation: on the complementarity of experimental and naturalistic studies" *Journal of American Society for Information Science and Technology* 52 (8):676-685.

Taylor, R.S. (1968) "Question-negotiation and information seeking in libraries" *College & Research Libraries* 29:178-194.

Tenopir, C. (2001) "Online databases: database and online system usage" *Library Journal* 126 (16):41-42, 45.

Toms, E.G. (1997) "Browsing digital information: examining the 'affordances' in the interaction of user and text." PhD dissertation. University of Western Ontario.

Toms, E.G. (2002) "Information interaction: providing a framework for information architecture" *Journal of the American Society for Information Science & Technology* 53 (10):855-862.

Vakkari, P., Savolainen, S. & Dervin, B. (eds.) (1996) *Information seeking in context: proceedings of an International Conference on Research in Information Needs, Seeking and Use in Different Contexts, 14-16 August, 1996, Tampere, Finland* (London: Taylor Graham).

Vickery, B. (1997) "Metatheory and information science" *Journal of Documentation* 53 (5):457-76.

Walker, S. & Jones, R.M. (1987) *Improving subject retrieval in online catalogues. 1, Stemming, automatic spelling correction and cross-reference tables* (London: British Library).

Wallace, P.M. (1993) "How do patrons search the online catalog when no one's looking? Transaction log analysis and implications for bibliographic instruction and design" *RQ* 33:239-252.

Wallace, P.M. (1997) "Periodical title searching in online catalogs" *Serials Review* 23 (3):27-35.

Wang, P. (1997) "The design of document retrieval systems for academic users: implications of studies on users' relevance criteria" in *ASIS '97: proceedings of the 60th ASIS Annual Meeting,* edited by C. Schwartz & M. Rorvig, pp. 162-173 (Medford: Information Today).

Wang, P. & Soergel, D. (1998) "A cognitive model of document use during a research project. Study I, Document selection" *Journal of the American Society for Information Science* 49 (2):115-133.

Wang, P. & White, M.D. (1995) "Document use during a research project: a longitudinal study" in *ASIS '95: proceedings of the 58th ASIS annual meeting,* edited by T. Kinney, pp. 181-188 (Medford: Information Today).

White, H.D. & McCain, K. (1997) "Visualization of literatures" *Annual Review of Information Science & Technology* 32:99-168.

Wiberley, S.E., Daugherty , R.A. & Danowski, J.A. (1995) "User persistence in displaying online catalog postings: LUIS" *Library Resources & Technical Services* 39:247-264.

Williamson, K. (1998) "Discovered by chance: the role of incidental information acquisition in an ecological model of information use" *Library & Information Science Research* 20 (1):23-40.

Wilson, T.D. (1999) "Models in information behaviour research" *Journal of Documentation* 55 (3):249-270.

Wyly, B.J. (1996) "From access points to materials: a transaction log analysis of access point value for online catalog users" *Library Resources & Technical Services* 40:211-236.

Xie, H. (1997) "Planned and situated aspects in interactive IR: patterns of user interactive intentions and information seeking strategies" in *ASIS '97: proceedings of the 60th ASIS Annual Meeting*, edited by C. Schwartz & M. Rorvig, pp. 101-110 (Medford: Information Today).

Xie, H. (2000) "Shifts of interactive intentions and information-seeking strategies in interactive information retrieval" *Journal of the American Society for Information Science* 51 (9):841-857.

Yee, M.M. (1991) "System design and cataloging meet the user: user interfaces to online public access catalogs" *Journal of the American Society for Information Science* 42 (2):78-98.