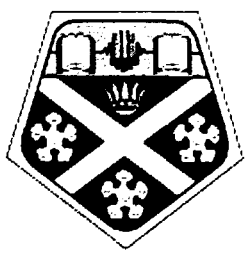


THE UNIVERSITY OF STRATHCLYDE IN GLASGOW

End-user Interaction with Thesaurus-enhanced Search Interfaces: An Evaluation of Search Term Selection for Query Expansion

Ali Asghar Shiri

Department of Computer and Information Sciences

University of Strathclyde

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

November 2003

Declaration of Author's Right

Abstract

A major challenge faced by end-users during the information search and retrieval process is the selection of search terms for query formulation and expansion. Thesauri are recognised as one source of search terms with the potential to assist users in the process of term selection. Research in search term selection, query expansion and interface evaluation has stressed the importance of providing end-users with terminological assistance. As the number of thesauri attached to information retrieval systems has grown, a range of interface facilities and features have been developed to aid users in formulating their queries.

This study investigated end-user interaction with a thesaurus-enhanced search interface to evaluate their search term selection and query expansion behaviour. The main objectives of this study were: to evaluate how and to what extent a thesaurus-enhanced search interface assisted end-users in selecting search terms for query expansion, to ascertain users' attitude toward both the thesaurus and interface as tools for facilitating search term selection, and to identify searching and browsing behaviours of users interacting with a thesaurus-enhanced interface. The test environment involved the Ovid CAB Abstracts database, the CAB thesaurus, and 30 academic staff and postgraduate students with genuine search requests. The data gathering tools employed were pre-search questionnaires, screen capturing software, post-search questionnaires, and post-session interviews.

The results demonstrated different patterns of thesaurus-based search term selection by academic staff and postgraduates. Academic staff with more extensive domain knowledge tended to select narrower terms whereas postgraduates more often chose related and broader terms. In general, all users selected a larger number of narrower and related terms for expanding their queries. The effect of topic characteristics such as topic complexity and topic familiarity on search behaviour was also investigated. It was shown that complex topics affected users' cognitive and physical moves, number of search terms selected and query expansion instances. Topic familiarity was also found to have an effect on users' browsing behaviour. An evaluation of users' perceptions of the interface indicated that usability was a factor affecting thesaurus browsing and navigating behaviour.

This study was constrained by the limitations of the IR system utilised, the experimental design and the choice of subjects. However, this study can be viewed as the first investigation of variables such as topic complexity and topic familiarity within a thesaurus-enhanced search environment. The findings of this study contribute to research in the areas of user-centred search term selection, thesaurus-assisted query expansion and the evaluation of user interaction with IR search interfaces.

Acknowledgements

This dissertation is dedicated to my father Zabihollah Shiri who passed away before its completion. His unquenchable thirst to read and learn will always remain a source of inspiration and motivation to me.

I am pleased to acknowledge all of the people who have contributed to my personal and professional development during my PhD research. First of all I would like to thank my supervisor Crawford Revie who has unstintingly supported and encouraged me during the course of my PhD. His kind and friendly personality was a great privilege as I could always discuss with him my personal concerns and problems. I am also grateful to my second supervisor Dr. Gobinda Chowdhury for his constructive comments and suggestions. I also wish to thank Professor Forbes Gibb who has been a valuable source of support from the very beginning of my studies in the department. I would also like to thank Professor Micheline Beaulieu, my external examiner, and Dr. Ian Ruthven for their useful comments during the viva.

I owe a debt of gratitude to the staff of the department, in particular, Monica, Linda, Zakia, Catriona and Duncan who have always been supportive, caring and kind to me.

Thanks to the University of Strathclyde for awarding me a postgraduate research student grant, providing me with financial support to complete my PhD research. I would also like to thank CAB International for providing me with unlimited access to their CAB Abstracts database which was used in my experiments.

I am truly indebted to my family Farzaneh and Kimia who have generously given me love and support through the ups and downs of my research during the last four years. I share this accomplishment with each and every one of my loved ones without whose support this thesis could not have come into existence.

Table of Contents

Chapter 3: Experimental Design and Methodology

Chapter 4: Descriptive Data Analysis

Chapter 7: Conclusions

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Research Rationale

The selection of search terms for query formulation and expansion is a key phase in the Information Retrieval (IR) process. Saracevic (1997) describes the search term selection process as a dynamic process which involves interaction between user and computer at various levels. The dynamic nature of this process implies that search terms are selected from various sources and adapted, changed or abandoned at various levels in the process.

Thesauri have been recognised as a useful source for enhancing search term selection for query formulation and expansion (Beaulieu, 1997; Brajnik et al., 1996; Fidel, 1991). Research on searching behaviour, IR interface evaluation, search term selection and query expansion has addressed the issue of providing users with terminological assistance to enhance information retrieval. Such assistance may be provided through the inclusion of thesauri and classification schemes within IR interfaces.

Researchers have investigated the searching behaviour of various types of users and have looked in particular at their search term selection

behaviour. These studies have suggested that the selection of terms can be improved if thesauri are incorporated into the search interface (Efthimiadis, 2000; Hsieh-Yee, 1993; Sutcliffe; 2000; Vakkari, 2000).

Another line of investigation, providing evidence for the benefit of incorporating thesauri into interfaces, deals with interactive query expansion and interface evaluation. In a series of experiments on designing interfaces to the Okapi search engine, it was found that both implicit and explicit use of a thesaurus during automatic and interactive query expansion was beneficial. It was also suggested that while the system can find useful thesaurus terms through automatic query expansion process, terms explicitly selected by users were of particular value (Beaulieu, 1997; Jones et al., 1995).

In addition to the above empirical studies, a number of interfaces enhanced with thesauri have been developed over the last decade (Agosti et al., 1992; Bates, 1990; Beaulieu, 1997; Belkin; 1993; Hearst and Karadi, 1997; John and Cochrane, 1995; Lin, 1999; Pollitt, 1994). However very few of these interfaces have been formally evaluated in terms of the ways in which they support query formulation and expansion. Nor have topic characteristics and users' perceptions of issues such as thesaurus interaction and usability in an operational environment been extensively addressed.

Recent developments in end-user searching and the wider availability of online information retrieval systems together with advances in user-centred interface design have led to the increased use of thesauri as search aids (Shiri et al; 2002). However, there has been only limited explanation of how users interact with these interfaces or of the ways in which end-users' search term selection for query formulation or expansion is affected.

The present study builds on previous research into search term selection and query expansion, thesauri as search term sources, information seeking behaviour, and information retrieval interaction. This study investigates end-users' search term selection and query expansion behaviour in a thesaurus-enhanced search environment. It is particularly concerned with the ways in which the thesaurus assists users in selecting search terms and in expanding queries. Specific issues associated with user-thesaurus interaction such as the identification and classification of moves made by users during the search process, users' perceptions towards the thesaurus and interface, and the effects on cognitive and physical move types of search topic characteristics have been examined.

1.2 Objectives and Research Questions

The main objectives of this study are:

- to evaluate how and to what extent a thesaurus-enhanced search interface can assist end-users in selecting search terms for query expansion.

- to ascertain users' attitude toward both the thesaurus and interface as tools for facilitating search term selection for query expansion.

- to identify searching and browsing behaviours of users while interacting with a thesaurus-enhanced interface attached to a large bibliographic database available on the Internet.

In order to address the above research objectives five key research questions were formulated.

- Are there common patterns of user behaviour in thesaurus-based browsing and searching?

- What relationships are there between users' initial query terms and the terms they select from the thesaurus for query expansion?

- Does topic complexity affect user-thesaurus interaction in general and search term selection in particular?

- Does topic familiarity affect user-thesaurus interaction?

- Does interface usability affect thesaurus browsing or other search behaviour?

1.3 Hypotheses and Sub-hypotheses

Based on the key research questions five main hypotheses were developed. A description of these hypotheses is provided below:

Main hypothesis 1: User characteristics have an effect on searching and browsing behaviour.

This hypothesis addresses users' personal characteristics and educational background and their effect on user interaction and search behaviour. The characteristics include users' subject knowledge, status, gender, and computer experience. User interaction and search behaviour are represented by cognitive and physical moves (defined in detail in the methodology chapter) and the number of terms browsed and selected.

Main hypothesis 2: Topic complexity affects user- thesaurus interaction.

This hypothesis is particularly associated with the complexity characteristic of search topics and its effect on users' interaction with thesauri. The interaction in this hypothesis was evaluated in terms of

cognitive and physical search moves, query reformulation and search term selection.

Main hypothesis 3: Topic familiarity and search type have an effect on searching and browsing behaviour.

The users' familiarity with search topics and their prior topic search experience are the main focus of this hypothesis. A set of sub-hypotheses were developed to evaluate the effect of topic familiarity, topic search experience and search type on cognitive and physical moves, number of search terms, and users' judgment on the value of additional thesaurus terms.

Main hypothesis 4: Search process characteristics affect search term and search result characteristics.

This hypothesis is concerned with the evaluation of search process characteristics such as cognitive and physical moves, search term selection and time and ways in which they affect users' judgement of result relevance, result number satisfaction and term usefulness.

Main hypothesis 5: Perceptions of interface usability have an effect on searching and browsing behaviour.

The last hypothesis investigates the relationships between thesaurus browsing/navigation and ease of using and learning the search interface. Variables include cognitive and physical moves, number of terms browsed and selected, number of query reformulation instances and time spent per search.

All five main hypotheses have been broken down into detailed sub-hypotheses and these are presented in the Methodology Chapter.

1.4 Original Contribution

The present study proposes an analytical framework for the quantitative evaluation of user interaction within a thesaurus-enhanced search interface. The framework is based on the identification of cognitive and physical search moves associated with individual search states in an end-user searching environment. User interaction with the thesaurus-enhanced search interface was examined to explore ways in which users perceive thesauri and their interfaces during search term selection and query expansion.

The present research may be considered to be the first study to have explored variables such as topic complexity and topic familiarity in an operational thesaurus-enhanced search environment, involving end-users searching for real information requests. It also investigates the interaction between these variables with respect to cognitive and physical moves. It sheds light on the ways in which cognitive and physical moves are associated with users' search term and search result satisfaction.

Another novel aspect of this research lies in its choice of test thesaurus. To date a limited range of thesauri have been utilised in user studies, predominantly the MeSH, INSPEC, ERIC, and ProQuest thesauri. This study chose to exploit the CAB thesaurus for test purposes. Although this thesaurus is known as one of the largest and most well-developed thesauri in the area of life sciences and agriculture, no study was found to have used the CAB thesaurus in an IR interaction environment.

Another contribution made by this investigation is its selection of an atypical research population. Most user studies to date have involved end-users from the fields of library and information science, computer science, social science, and medicine. This study involved academic staff and postgraduate researchers from an under-represented end-user

community, namely veterinarians and life scientists. This is important because it helps extend our knowledge of end-user behaviour from a wider range of educational backgrounds. The results of this study can be compared with those in other areas such as social sciences or engineering to establish whether or not the findings are generalisable in other domains.

The results obtained from this study made contributions in the following areas.

End-user searching: the study is important because it reveals issues and problems surrounding end-user interaction with thesauri within the broader context of the search process and the ways in which users' search and browse behaviour are affected.

Thesaurus-based query expansion research: the results help extend our knowledge of the types of search terms end-users select while interacting with a thesaurus-enhanced search interface. This knowledge can be used to improve thesaurus-based query expansion through increased availability of the most widely used types of thesaurus terms and relationships.

Thesaurus-enhanced interface design: knowledge about the ways in which end-users perceived and interacted with the interface together with the usability problems reported can be used to inform the effective incorporation of thesauri into search interfaces.

1.5 Structure of Thesis

This thesis is arranged into seven chapters, a reference list and accompanying appendices. Rationale for the research, objectives and research questions, hypotheses and sub-hypotheses, and original contribution have been introduced in Chapter 1. The purpose of Chapter

2 is to provide the context to and justification for the research reported in this thesis. It examines the research literature dealing with users' search behaviour with particular emphasis on search term selection. This chapter is divided into three sections: information seeking and information retrieval interaction models, thesaurus-assisted search term selection and query expansion from a user-centred perspective, and a review of thesaurus-enhanced search interfaces. Chapter 3 provides the general theoretical and evaluation framework underlying this research. It also discusses the research design and experimental methods, tools and techniques which formed the environment in which this research was conducted. Chapter 4 presents the descriptive findings from the experiment reported in Chapter 3. The descriptive findings are based on six sets of variables, namely: user characteristics; search topic characteristics; search process characteristics; search term characteristics; user satisfaction characteristics; and usability characteristics. In Chapter 5 results of the statistical data analysis and hypotheses testing associated with the experiment are reported. The purpose of Chapter 6 is to discuss and suggest interpretations of the descriptive and statistical findings reported in Chapters 4 and 5. The conclusions drawn based on the discussion and interpretations are presented in Chapter 7 which also discusses the limitations and implications of the research together with recommendations for further research.

Chapter 2: Literature Review

The aim of this chapter is to provide the context and background to and the justification for the research reported in this thesis. It highlights the main issues surrounding the evaluation of user searching behaviour, particularly the selection of search terms for query formulation and expansion using thesauri. The chapter is divided into three main sections: information seeking and information retrieval (IR) interaction models, thesaurus-assisted search term selection and query expansion from a user-centred perspective, and a review of thesaurus-enhanced search interfaces.

2.1 Information Seeking and IR Interaction Models

This section provides a review of information seeking and IR interaction models and their implications for search term selection, query formulation and expansion. Particular focus is given to an assessment of the extent to which these models have considered the subject approach and in particular any thesaurus component within their modelling of the query formulation and expansion process.

The role of any model is to depict the main elements of and relations between objects (system, process, entity, structure, idea). Scientific models are characterised by the property that they are testable. Therefore, models themselves are a subject of examination and critique in terms of how well they depict an object (Saracevic, 1997). Models in the general field of information behaviour may be described as a framework for thinking about a problem: they are statements, often in the form of diagrams, that attempt to describe an information-seeking activity, the causes and the consequences of that activity, or the relationships among stages in information seeking behaviour (Wilson, 1999).

The development of most of the models for IR interaction can be traced back to the 1980s and 1990s. During this period, a number of models were proposed to address the issues related to understanding the complex nature of human-information interaction. These models have approached the issue from a range of perspectives. IR interaction, information-seeking behaviour, information searching behaviour and information behaviour were among the terms used to denote the features of these models. As Beaulieu (2000) has commented, information seeking, information searching, and IR interaction have traditionally been the concerns of separate research communities having the interaction element in common. Taken together they provide complementary views of a highly dynamic process.

In a review of information behaviour models, Wilson (1999) identifies two major types of models: Information seeking behaviour models and Information searching models. He looks into these models from the information behaviour perspective. He considers this perspective to involve those activities a person may engage in when identifying his or

her own information needs, searching for such information in any way, and using or transferring that information. This perspective is more general than that of most IR researchers who view both information searching models and information seeking behaviour models as its sub-sets. To examine individual models related to information seeking and searching, and IR, Wilson's categorisation is used here to divide the models into two groups: models of information seeking behaviour and IR interaction models.

2.1.1 Information Seeking Behaviour Models

Models of information seeking behaviour try to depict all steps, stages and efforts undertaken by users in interacting with information from a wide range of information sources regardless of their formats, coverage, and presentation. This implies all situations, tasks and activities users deal with from the instantiation of the information needs to information use and satisfaction. As these models include generalised descriptions of information seeking behaviours, only models which make explicit reference to query formulation and the subject approach are examined here.

Studying the information seeking patterns of academic social scientists, Ellis (1989) derives a behavioural model with implications for IR system design. The model includes six characteristics, which constitute the principal generic features of the different individual patterns. These features are named as *starting, chaining, browsing, differentiating, monitoring, and extracting.* Based on this model and a detailed analysis of the social scientists' behaviour, he found subject access i.e. browsing subject terms to be a general problem in browsing and recommended features for IR system design. The browsing feature in his model has significant bearing on the subject approach and the use of thesauri for

search term selection and query formulation. He suggests that some form of broader and narrower subject description is necessary to provide browsing facilities for exploring a retrieval system. His model suggests that it would be useful to provide searchers with the possibility of browsing the hierarchical structure of a thesaurus to broaden or narrow their search as well as when looking for related terms for inclusion in the search formulation.

Bates's (1989) model of browsing and berrypicking is one of the most well-cited models of the search process. She has used the term berrypicking by analogy to picking huckleberries or blueberries in the forest in order to develop a model of evolving query formulation and modification. The model touches upon different behavioural, heuristic and physical aspects of the process that searchers might encounter while they are involved in online searching. Query and search terms in this model are perceived as dynamic entities which vary in part or whole to satisfy an information need throughout the search process. Considering manual information seeking behaviours, she tries to highlight various capabilities that users might be interested to have during the search process. She refers to six main strategies users employ: footnote chasing, citation searching, journal run, area scanning, subject searches in bibliographies and abstracting/ indexing services, and author searching. The notion of area scanning refers to browsing the materials that are physically collocated with materials located earlier in a search.

Two of these strategies, namely *area scanning* and *subject searches in bibliographies and abstracting/indexing services,* are strongly related to the subject approach to the information search process. In the first strategy, the subject approach to library catalogues is the focus, while in the second subject indexes and categories are referred to as tools for

searching the information. Among the design features Bates suggests based on her berrypicking model is the provision of capabilities for users to browse general as well as subdivisions of the classification categories used in an abstracting and indexing service while searching.

Kulthau's (1991) model of the information search process has been derived from a series of five studies investigating common experiences of users in information seeking situations. Stages of the information search process identified in the model are *initiation, selection, exploration, formulation, collection*, and *presentation.* The model depicts the information seeking process from the users' perspective. The selection and formulation stages are both similar in nature to the elements of user-system interaction. In this model, the selection stage mainly deals with topic selection with an overview of alternative topics which can be compared to the selection of search terms in the interaction process. The formulation stage is seen to be an important aspect of the search process as the user tries to form a focused perspective of the topic. This stage can also be compared with the formulation of a proper query to find relevant information. These two stages have implications for query formulation and reformulation during the search process.

The models discussed so far attempt to model human information interaction from a broad perspective covering the universe of needs, behaviours, strategies, and information sources (in a very general sense) and are strongly influenced by social, psychological, and communication theories. The next section provides an overview of IR interaction models and the extent to which they have considered the use of thesauri in the search term selection process.

2.1.2 IR Interaction Models

Models classed as IR interaction models focus on the nature of the interaction between users and computerised IR systems. The rationale behind these models is to provide a holistic view of all components, processes, actions and steps involved in user interaction with an IR system.

A number of IR interaction models have been proposed, some of which have undergone various developmental stages to take into account as many human-system interaction dimensions and perspectives as possible. Here, an attempt is made to give an account of the major models that have implications for the present investigation.

Ingwersen's (1996) *cognitive model* can be viewed as one of the major IR interaction models that concentrates on different cognitive structures affecting IR interaction. This model is an extension and completed version of the previous cognitive models proposed by Ingwersen (1982,1992). The model implies that human cognitive structures are inherent in various stages of IR interaction. There are elements that can affect the user cognitive state while interacting with an IR system which Ingwersen refers to as 'cognitive origins'. The user and his cognitive space lie in the heart of the model and different cognitive transformations occur between and among the elements depicted in the model e.g. user interacting with intermediary, intermediary interacting with the IR system, and so on. He states that a variety of human actors including system designers and producers, indexing rule constructors, indexers, authors of texts and images, intermediary mechanism designers, and users all possess cognitive structures which can influence the IR interaction. For example, index terms are representations of the human indexers' cognitive structures added to the original information objects which in

turn affect the formulation of the query. Thesauri can be construed as cognitive structures which interact with other cognitive structures and influence the IR interaction in general and query characteristics in particular. Ingwersen also suggests the notion of polyrepresentation to refer to different cognitive structures such as problem statement, information need, and current cognitive state. From the polyrepresentation point of view, thesauri are viewed as one type of representation of information objects, which coupled with other types of representation, affect the users' decision making and interaction. User-thesaurus interaction in this model can be construed as the interaction between cognitive constructs of users and indexers where the semantic structure of the thesaurus influences the users' articulation of an information need and well-defined query construction.

The *Episode model* of Belkin et al. (1993b, 1995) is another IR interaction model which attempts to consider user interaction with *texts* (any information-bearing objects) as the central phenomenon of IR. The central process for IR in this model is the interaction of user with text, in support of a wide range of information seeking behaviours and related goals, relevant to different problematic situations. Each information seeking strategy can be defined as an *episode*. Belkin points to the useful distinction between *searching* for a given item and *scanning* to find interesting items within the information collection. The dimension of "goal" in the interaction implies the intention of users in dealing with an IR system which may be *learning* about some aspect of the resource or *selecting* useful items for retrieval. The mode of retrieval signifies the extent of the user's knowledge about the location of the item. If s/he looks for identified items the model of retrieval would be called retrieval by *specification* as the item has already been specified. If the user does not have a clear idea of the item and through scanning similar items finds the item, the retrieval mode is referred to as retrieval by *recognition*.

Finally, the fourth dimension of the model refers to the two types of information resource users interact with. *Information* which is the item through which the user's need can be met. *Meta-information* sources such as thesauri and classification systems that may be consulted by users to characterise how information resources have been represented in the IR system. In this model, the use of thesauri is interpreted as one type of information seeking strategy which can be used in at least eight of the sixteen strategies proposed. Interaction of users with thesauri is seen as related to different searching and browsing strategies.

The *stratified model* of IR interaction proposed by Saracevic (1997), attempts to cast light on the complex human-system interaction process and its constituent elements. IR interaction in this model is conceived as a layered dialogue between the participants - user and computer - through an interface, the main purpose of which is to affect the cognitive state of the user to ensure more effective use of information in connection with the task at hand.

From the user interaction perspective, the model depicts such processes as search term selection from different sources for query specification and modification as well as search tactics which occur at the interface level. All these processes are considered to be influenced by a range of cognitive, affective and situational characteristics of the user. At the cognitive level users interact with texts and their representations in the information sources. Users interact with their intentions and all that go with intentionality such as beliefs, motivation and feelings at the affective level. At the situational level users interact with a given situation or problem-at-hand which produced the information need. Query characteristics are considered to be an important component of the user dimension of the model.

The model views search term selection as a dynamic interactive process within IR, which is subject to a range of changes and shifts throughout the search process. One of the advantages of Saracevic's model is its extension (Spink and Saracevic, 1997) for search term selection which provides some useful implications as to how the model can be used in an empirical environment. Within this model, thesauri, index terms, classification schedules together with other term selection sources are viewed as possible sources of term selection during interaction.

2.1.3 Conclusion

A number of information seeking and information searching models have been introduced to explore and identify those interaction elements and processes which are relevant to research in thesaurus-aided search term selection for query formulation and expansion. The information seeking models provided some general recommendations to include subject tools such as thesauri and classification schedules in retrieval systems to support the information seeking process. In contrast, more specific implications were found in IR interaction models for search term selection, query formulation and the role of thesauri. This was due to the fact that these models depict the specific interaction of users with computerised retrieval systems. Ingwersen's model (1996) conceptualised thesauri, index terms and users as cognitive structures which constitute the interaction elements and affect the whole IR process. Thesauri were perceived as supportive tools for browsing and searching within the episodic model of information seeking strategy (Belkin et al., 1993b). Query characteristics, search term selection and the sources of search terms (such as thesauri) as well as shifts and changes in the query formulation process were elucidated in the stratified interaction model (Saracevic, 1997).

While all of the reviewed models make explicit or implicit reference to thesauri and their role in IR interaction in general and search term selection in particular, the ways in which users interact with thesauri and the extent to which this interaction affects their term selection behaviours calls for more research. The present study investigates detailed aspects of user interaction with thesauri and the ways in which this affects search term selection for query formulation and expansion.

2.2 Thesaurus-assisted Search Term Selection and Query Expansion: A Review of User-centred Studies

This section provides a review of the literature related to the application of domain-specific thesauri in the search and retrieval process. Focusing on studies that adopt a user-centred approach, the review presents a survey of the methodologies and results from empirical studies undertaken on the use of thesauri as sources of term selection for query formulation and expansion during the search process. It summarises the ways in which domain-specific thesauri from different disciplines have been used by various types of users and how these tools aid users in the selection of search terms. The review consists of two main sections: first, studies on thesaurus-aided search term selection; and second, studies dealing with query expansion using thesauri. Both sections are illustrated with case studies that have adopted a user-centred approach. (An earlier version of the literature reviewed in this section appeared in Shiri et al., 2002a).

2.2.1 Introduction

The selection of search terms for query formulation and expansion is a challenging task within the information search and retrieval process. Two general approaches have been adopted in studies on search term selection: system-centred and user-centred. The system-centred approach is represented by work on algorithms and evaluation based on

the traditional IR model, a model that fundamentally ignores the users and their interaction with the system. In contrast, the user-centred approach focuses on the cognitive, interactive, and contextual aspects of IR and considers users, use, situations, context, and interaction with system (Saracevic, 1999). The user-centred approach has been developed to address a range of poorly understood issues relating to behavioural and cognitive aspects of the IR process. Spink (1994a) and Spink and Saracevic (1997) have emphasised the need for further research into the user-centred approach to search term selection and query expansion in order to improve the use of and interaction with IR systems. This approach is concerned with the ways in which users of IR systems select their search terms for formulating and/or expanding their queries. It also considers factors and variables that cognitively and behaviourally affect the user's decision-making in the search term selection process.

Knowledge structures in general and domain-specific thesauri (also referred to by some as subject-specific thesauri) in particular are potential sources of search terms for query formulation and expansion. Several studies have evaluated the use of thesauri by different types of users. While all of these studies have adopted user-oriented approaches, their treatment of the use of thesauri, types of users and methodologies varies.

The main objective of this review is to provide a survey of the methodological issues and main findings of a relatively comprehensive set of user-centred studies on thesaurus-aided search term selection and query expansion. Two criteria were taken into consideration when defining the scope of this review. First, the review focuses on studies that have considered thesauri as sources of term selection and query

expansion. Second, since there are different types of thesauri, only studies that have applied or evaluated domain-specific manually constructed thesauri with standard relationships (hierarchical, equivalence and associative) are included in this review. Thus, for example, studies that have evaluated the role of automatically constructed thesauri in search term selection and query expansion are excluded. The reason for choosing a manually constructed thesaurus for the present research is due to the fact that this type of thesaurus represents semantic relationships defined by humans and involves an intellectually reliable set of terms and relationships. The second reason is based on previous research (Fidel, 1991b; Beaulieu, 1997; Greenberg, 2001b) that demonstrated that manually constructed thesauri were found to be beneficial to the search term selection and query expansion process.

This review is structured into two major sections. The first section deals with the use of thesauri for search term selection. The second section is concerned with the application of thesauri for query expansion purposes to improve search results. The main issues of each section are summarised and an overall conclusion ends this section.

2.2.2 Approaches to Search Term Selection

The selection of search terms for query formulation and expansion within the IR process has been studied from a range of perspectives. Spink and Saracevic (1997) identified two general types of search term selection research, namely the algorithmic and human approaches.

The focus of the algorithmic approach is to develop and evaluate different types of algorithms for selecting, weighting and/or ranking search terms in the process of query formulation or expansion to improve information

retrieval. Examples of research of this type include (Spark Jones, 1979; Van Rijsbergen et al., 1981; Salton and Buckley, 1988; Robertson, 1990; Efthimiadis, 1993; Robertson et al., 1997; Magennis and Rijsbergen, 1997) as well as much of the research documented in the Text REtrieval Conference (TREC) proceedings.

The human approach, in contrast, is concerned with studying and evaluating the ways in which users choose terms for formulating, expanding or modifying their queries during the search process. It deals with cognitive and behavioural models and issues that affect the selection of search terms by users. Research has focused on user-centred variables such as those relating to information needs, user intentions, personal characteristics, and different user information seeking profiles, and investigates their relationship to term selection in the search process.

2.2.3 Sources of Search Term Selection

A number of sources for search term selection in query formulation and expansion have been suggested in the literature. For instance, Fidel (1991a) analyses search term selection based on two types of source: free-text and controlled vocabulary descriptors. Efthimiadis (1996) categorises sources for term selection in query expansion into two types: those based on the relevance feedback process to initial search results and those that use some form of knowledge structure. These structures can be either collection-dependent (corpus based) or collection-independent such as thesauri and dictionaries. Spink and Saracevic (1997) identify five sources of search terms in a study investigating their effectiveness during mediated online searching. These sources are: the question statement, user interaction with the intermediary, thesaurus, the human intermediary, and term relevance feedback. In these studies and those reviewed in the following sections thesauri were recognised as

tools suited to the provision of search terms at either the query formulation or expansion phases of the IR process.

2.2.4 Thesaurus-aided Search Term Selection: User Studies

Studies within this area have investigated information searching behaviours, tactics and strategies to better understand term selection and the ways in which thesauri are utilised by users during this phase of the search process.

Bates (1979) in a discussion on search strategy, defines four categories of search tactics, two of which are relevant in the present context; these she labels search formulation tactics and term tactics. Search formulation tactics refer to the process of designing or redesigning the search formulation and ways of analysing information request and query elements. This includes making the search formulation precise, broad or specific. Term tactics relate to the selection and revision of specific terms within the context of search formulation. These tactics focus on moving upward, downward, or sideways within a hierarchical structural to find broader or narrower terms during the selection of search terms. These tactics attempt to capture the complexity of the human search term selection process and the sophisticated decision-making effort involved.

Several studies have investigated search term strategies and the term selection behaviour of users from a wide variety of disciplines, backgrounds and environments. Although treated differently, thesauri were considered or evaluated as search term sources in all of the studies discussed below. These studies can be categorised according to the user population involved:

- professional searchers only

- professional searchers and end-users in a mediated environment
- professional and novice search behaviour compared; and
- end-users only

Search Term Selection Behaviour of Professional Searchers

While investigating the searching behaviour of professional online searchers is a well-developed research theme, a number of key studies concerned with the use of thesauri are due to Fidel. In an early investigation Fidel (1984) studied the search formulation, reformulation, and search term selection behaviour of five experienced searchers. She identified two types of information searching behaviours which she named the *operationalist* and *conceptualist* searching styles. Operatioanlist searchers try to formulate a query by identifying the related descriptors for each component of the request. They look for the descriptors and check their categories and locations in the hierarchical structure not only to find permitted entry terms but also to gain a better understanding of the request. In situations where they cannot find a descriptor to represent a concept they search using the free-text mode. In contrast, conceptualist searchers focus on the structure of the vocabulary for conceptual analysis of the request and query formulation. They look for relationships between the facets of the request and the structure of the controlled vocabulary. They tend to use free-text terms for the initial search formulation in cases where they are confident that there is no controlled way to express the concept, or where terms are very specific and well-defined, and very little has been written about the subject. Both searching styles represent the term selection styles of professional searchers using thesauri and the ways in which they cognitively process terms in the request and translate them into queries accepted by the retrieval system.

In another study Fidel (1986) observed the online searching behaviour of eight searchers in order to provide a set of rules for search term selection by an intermediary expert system. The search keys selected by the searchers and the reasons for choosing these terms were examined. Using this evidence a "selection routine" detailing the conditions for the selection of search keys from both free-text and controlled vocabularies was developed. The selection routine describes different conditions under which searchers try to map a search term to a descriptor in the thesaurus or controlled vocabulary, and those which lead to a decision to use the term as a free-text key. Mapping to descriptors can be carried out through an exact match, a partial match, or the key might be mapped to a broader descriptor or to narrower terms. By analysing the patterns of selection routine, Fidel illustrated the significance of decisions made during search key selection to the success of a search. Her model also showed that a "good" search term is a single-meaning term that can be mapped to a descriptor, while a term is "not adequate" if it is a common term and/or cannot be mapped to a descriptor. She refers to the lack of a general typology for requests to support the selection of different ways of choosing terms and suggests that more investigation should be undertaken on searching behaviour to reveal under which conditions each of the term selection options is chosen. The "selection routine" approach identifies the problematic points in the search term selection process and provides guidelines for research into the searching behaviour of online searchers.

In a larger investigation of the search term selection process, Fidel (1991a) studied 47 professional online searchers performing job-related searches. Two distinct types of search keys: words used in free-text searching, and descriptors taken from a controlled vocabulary, were compared. In the first part of this study she built on her previous work (1986) on the validation and expansion of a search term selection routine.

She defined *single-meaning terms,* as those which are good for free-text searching, and *common terms* for those which are not. In order to build a formal model of search term selection, different conditions of selecting a term were identified. For example, whether a common or single-meaning term can or cannot be mapped to a descriptor and, if it can be mapped, is this an exact or partial match. Different strategies for choosing broader, narrower or synonymous terms were identified. Searchers who participated in the study were asked to provide reasons for the selection of search keys. These reasons were categorised as: request-related, database-related and searcher-related. The first dealt with the characteristics of the requests, the second was associated with facilities and characteristics of the databases used by the searchers (for example, their use of thesauri or other features), and the third related to the individual searching behaviours and habits of the searchers. These factors influenced the way in which searchers chose free text or thesaurus descriptors. For instance, if a request were very specific, the searcher would tend to use free text terms; or, if a database did not have a thesaurus, the searcher would be more inclined to enter free text terms.

In the second part of the study Fidel (1991b) observed how the 47 professional searchers used descriptors and text words as search terms. A number of variables were defined to measure the factors affecting the selection of search keys, including: institutional setting, subject area, databases used, number of search keys, number of moves, ratio of free-text keys to total number of keys, and *thesaurus neglect* ratio. The results showed that of the 3,200 search keys selected 50 percent were descriptors and 50 percent were free-text terms. In an analysis of all the search terms used it was revealed that searchers consulted a thesaurus for 75 percent of the case selections. Characteristics of databases, thesauri and requests were the major factors affecting search key selection. One of the main results of this part of the study was that

searchers used a thesaurus when it was of satisfactory quality and was easily available to them. In contrast, the non-availability of a thesaurus tended to increase the number of search keys and the number of moves in a search, and thus increased the effort necessary to perform a search. Fidel also suggested that database designers and search-system vendors should encourage the use of thesauri by designing easy-to-use and flexible thesauri, particularly as reliable sources of synonyms.

The third part of Fidel's study (1991c) explored the searching styles of the professional searchers specifically in relation to their modification of search strategies. She identified different moves for increasing and reducing the size of retrieved set and moves to increase both recall and precision with regard to the *operationalist* and *conceptualist* searching models (Fidel, 1984). She stated that the association between the number of search keys and the total number of moves was a significant pattern in online searching behaviour. It was revealed that the searchers who made more moves were likely to use more search keys than were searchers making fewer moves. Searchers selected, on average, around 13 keys per search, but the average number of keys varied greatly between searcher, ranging from just under three to almost seventy. The results also demonstrated that the average number of keys per search is typical for a searcher, and that a person's searching style will thus determine how extensive will be their use of terms. It was found that searchers considered recall to be the most important factor when they select search keys and also when they modify search strategies.

Search Term Selection Behaviour of Professional Searchers and End-users in a Mediated Environment

Studies considered in this section examine the search term selection and interaction behaviour of professional searchers and end-users in a mediated retrieval environment.

In a large-scale empirical investigation to characterise the elements involved in information seeking and retrieving, particularly from the cognitive and human decision making perspective, Saracevic et al. (1988) examined the selection of search terms by different searchers for the same questions. User, question, searcher, search, and items retrieved were the main variables observed in this study. Forty users, thirty-nine searchers and forty questions constituted the experimental environment, and different techniques such as questionnaires, interviews, transaction logs and videotape were used for data collection. This study made a significant effort to explore the different cognitive structures involved in the IR interaction process, by examining the following: users, intermediaries (searchers), IR systems, questions and their interplay. In order to analyse and evaluate all variables influencing the search and retrieval process, the study defined four sources of search terms: a) terms derived from an oral statement about the problem recorded by the user, but without any reference to the written question; b) terms extracted from the recorded oral statement and the written question submitted by the user; c) terms from the written question using only the words in the question as search terms without any further elaboration and; d) terms derived from written question plus terms from an appropriate thesaurus for elaboration. The results indicated that searches based on the user's written question plus the use of a thesaurus were rated as the second best searches in terms of recall and precision, outperformed only by the relatively intensive approach that used oral and written user statements. The research demonstrated that given the same question, different

searchers tend to select a few common terms and a considerable number of terms that are different. Searches based on different search term selection sources produced a significant difference in recall but no significant difference in precision. While end-users were involved in the research, this study examined primarily the process of search term selection by professional searchers and the ways in which they evaluated the search requests and selected terms based on the four term sources noted above.

In another study by Spink and Saracevic (1997) investigating the selection and effectiveness of search terms, the searches were again performed by professional searchers but end-users played an active role in selecting search terms during online sessions. The data consisted of the interactions of 40 faculty and doctoral students with 4 professional searchers, searching 40 questions provided by the participants. Interviews, questionnaires, and transaction logs together with video recording were the main data gathering techniques. The variables defined in the study were as follows: user satisfaction ratings, search outcome variables (including, number of relevant and non-relevant items retrieved and precision measures), search process variables (including, number of cycles and moves), and user characteristics such as domain knowledge. They identified and classified sources of search terms as follows:

Question statements: search terms derived from the user's written request

User interaction: search terms suggested by the user prior to and/or during the online search, but not included in the user's question statement

Thesaurus: search terms derived from a thesaurus associated with the database

Intermediary: search terms suggested by the search professional prior to and/or during the online search

Term relevance feedback: search terms suggested either by the user or professional searcher taken from the retrieved items identified by the user as relevant

They evaluated the effectiveness of each source and their contribution to the search results. Question statements and interaction with the user were responsible for 38 percent and 23 percent of the selected terms respectively, with thesauri contributing an additional 19 percent of the terms. A further 11 percent of the search terms came from term relevance feedback, while professional searchers were responsible for the remaining 9 percent. In addition to supplying the largest proportion of terms question statements were also the most productive in terms of retrieving relevant items. User-interaction terms were slightly less effective with around 50% resulting in relevant retrieved items. Terms derived from thesauri were less effective again, a fact which caused Spink and Saracevic to conclude that thesaurus terms prove most effective when combined in search statements with user terms. This finding emphasises the significance of interaction between users' terms and the terms taken from the thesaurus. Although thesauri were one of the sources of terms used during the search process, the interaction between the end-user and thesauri was not examined in this study.

In addition to the above studies, there are a number of investigations that have explored the interaction between users, professional searchers and IR systems in a mediated environment. These studies have considered the process of search term selection in less detail, and have not specifically examined thesauri as sources of term selection (Spink, 1996; Spink et al., 1996; Spink et al., 2002)

Search Term Selection Behaviour of Professional and Novice Searchers Compared

To compare the search term selection and information searching behaviour of users with different levels of experience, a number of studies have employed professional and novice searchers. In an investigation of the behaviours associated with the process of online bibliographic searching, Fenichel (1981) examined the differences among users with different levels of experience, while they were searching online systems. Five groups of searchers with different levels of online searching experience performed searches using ONTAP, a subset of the ERIC database. The main variables studied included: environmental, searcher, search process, and search outcome factors. Variables associated with the search process were: the number of commands used, free-text and thesaurus terms chosen, sets viewed, cycles, search modifications, and connect time. The results showed that all five groups used more thesaurus terms than free-text terms. The most experienced group used a significantly higher proportion of descriptors taken from the thesaurus than did the other groups. In addition searchers with ERIC experience used significantly more thesaurus terms than subjects without such experience. The study showed that having experience of databases equipped with thesauri affects the ways in which searcher select terms. However, the finding that novice searchers also make use of thesauri indicates the importance of recognising that thesauri can be useful sources of search terms for users with varying levels of experience.

Hsieh-Yee (1993) investigated the effects of subject knowledge and search experience on novice and experienced searchers' use of search tactics in online searches. Using transaction logs, on-site observation and think aloud techniques, she studied the online searching process of 33 professional searchers and 30 novice searchers. Based on previous search tactics, she defined a number of variables related to search term

selection, such as use of searcher's own terms, the searcher's reliance on the thesaurus structure, off-line term selection efforts, online usage of search terms, inclusion of similar concepts and synonyms, and the searcher's combination of terms. To evaluate the effect of subject knowledge the study defined two familiar and two unfamiliar questions for both novice and professional searchers to consider. The results showed that when searching a topic of which searchers had some knowledge, experienced searchers included more synonyms and tried more combinations of search terms than did novice searchers. Experienced searchers looked to the thesaurus for term suggestion and tended to formulate a more comprehensive search, while novice searchers tended to rely on their own terms. Novice searchers consulted the thesaurus much less frequently than did experienced searchers. While the use of thesauri as a term selection tactic was considered in this study, the interaction of searchers with the thesaurus interface was not examined.

In a study of differences in search term selection between the most and the least consistent searchers, Iivonen (1995) evaluated the inconsistencies among 32 subjects, 24 experienced searchers and 8 undergraduate students of information studies. The subjects were given the option of using both descriptors from a thesaurus and free-text terms to perform specified search tasks. Three factors relating to the terminological style of the searchers were identified: the number of search terms per search request, the number of search terms per search concept, and the proportion of descriptors among search terms. The results suggested that as the number of search terms increases the term consistency is adversely affected. However, increasing the proportion of descriptors will lead to improved term consistency. Those searchers who chose only a few search terms per request and per search concept, and attempted to ensure that these terms were controlled vocabulary descriptors, achieved higher inter-searcher consistency than the average.

These results are in line with prior expectations in that, by selecting from the controlled vocabulary the number of potential search terms is already limited. It was also shown that differences in the searchers' experience resulted in the use of different terminological styles.

Based on an empirical study of searchers during the pre-online stage, Iivonen and Sonnewald (1998) proposed a cognitive model of the search term selection process. Once again their study population was made up of 24 professional searchers, who had backgrounds in special, university and public libraries, together with eight students. Each searcher was presented with twelve requests to formulate query statements prior to the search process. Searchers used a Finnish database to select descriptors but also had the option to incorporate free-text terms in formulating their queries. The results of the study revealed six different sources of term selection, referred to by the researchers as six *discourses*. By discourse they meant a specific way of thinking and talking about a certain topic within a community. These six discourses, which constituted the elements of their proposed model, were: controlled vocabularies, documents and domain, indexing practice, client's search request, databases, and the prior search experience of the searcher. Documents and domain refer to the titles and abstracts of records, which represent the way the topic is discussed within a community of authors and publishers. Indexing practice relates to a searcher's perception of indexing rules and the practices adopted by indexers in the use of particular terms and concepts. Database discourse refers generally to the content and structure of a database and specifically implies knowledge of the subject categories and fields available in a given database. The study revealed that all but two of the searchers used controlled vocabularies. Irrespective of the searcher's background the only discourse which was frequently cited as a source of search terms was the controlled vocabulary. This illustrates the strong influence which the discourse of controlled

vocabularies exerts upon professional searchers. Controlled vocabularies provide a mechanism to describe a topic and to aid navigation by showing the relationships between topics. The results of this study illustrated the multidimensionality and complexity of the search term selection process and provided insights into the ways in which searchers navigate different discourses. The study suggested that further research was needed to fully identify and validate the characteristics of each discourse and to explore those aspects which facilitate the search term selection process.

Search Term Selection Behaviour of End-users Searching Independently

While there have been a number of studies on end-user searching behaviour, few studies have focused on search term selection.

Among the studies which consider discipline-oriented search terminology are the work of Bates et al. (1993) and Siegfried et al. (1993) which provide insights into the search terms used by humanities scholars as end-users. Their research examined the search techniques, queries and search terms of 27 scholars searching DIALOG databases. The researchers defined three major terminological categories in order to identify the vocabulary used by humanities researchers, namely: type of search need, such as works of an author or works on a subject; bibliographic features, such as date or form of publication; and types of subject, including individuals, geographic locations, date or period. While there were cases of using controlled vocabulary descriptors during search formulation, the study did not examine how and to what extent the searchers chose such descriptors or their impact on search term use. The results of the study showed that the terminology used by humanities researchers was remarkably different from the vocabulary used in other fields, as were aspects of the information seeking and online searching behaviour. The humanities scholars searched for more named

individuals, geographical terms, chronological terms and discipline terms. This finding has significant implications for developing thesauri and online search aids for the humanities, suggesting that thesauri developed for the humanities should incorporate more comprehensive sets of geographic and chronological terms as well as proper names.

Sutcliffe et al. (2000a) investigated end-user behaviour by studying the performance of 17 medical students searching the Medline database. The participants were categorised as expert or novice searchers based on their knowledge of the search system. Search performance, query pattern, search strategy, query construction, term use, and system facility usage were the variables studied. The study found that more than 80 percent of the expert searchers used the thesaurus or term suggestion facilitates to explore concepts. However no reference is made as to how the participants interacted with the thesaurus or what difference this made to search term selection. Expert searchers used more terms in their queries, constructed more complex queries, performed more iterations and used more system facilities. Novices, in contrast, used simple queries and fewer search terms. They made less use of system facilities and carried out fewer search iterations. The research concluded that although there were behavioural differences between novice and expert searchers, no simple correlations between behaviour and performance were found. A number of usability problems with the system thesaurus and term suggestion facilities were found. Suggestions were made as to how access to alternative terms might be enhanced and how the effort required to use the thesaurus or other term suggestion facilities reduced.

Vakkari (2000, 2001) investigated changes in search term usage and tactics among eleven students carrying out a research proposal as part of their Master's theses. Using survey questionnaires, interviews, the think-

aloud technique and transaction logs, he studied the search behaviour of students as they performed this task. He examined the number and types of search terms and tactics used and their relationship to the students' prior knowledge of their topics. He analysed different types of new terms introduced by the students and classified them as broader, narrower, related and synonymous terms. The results demonstrated that a growing focus and clearer understanding of the task led students to choose narrower and synonymous terms, to discard broader terms from their search formulation, and to use simpler search tactics. He suggested that for novices in a domain, structured terminological support would not only improve search results by encouraging the use of an increased number of narrower terms, but might also support the user in deconstructing the topic and interrelating its constituent parts.

A number of studies on end-user behaviour in different environments have commented on the specific issue of term selection with regard to IR interfaces (Marchionini et al., 1991; Meadow et al., 1995; Brajnik et al., 1996). Belkin et al. (2001) have reported the use and effectiveness of term suggestion facilities for supporting end-users in an interactive and relevance feedback environment. Another line of investigation adopted by researches has been the study of search term selection patterns used by children as part of their information seeking and retrieval behaviour (Marchionini, 1989; Solomon, 1993). While these studies have not evaluated thesauri as sources of term selection as such, their results indicate the need to incorporate thesauri and terminological support in the IR interface to facilitate query formulation by end-users.

In addition to the above studies, research examining search term use among end-users of World Wide Web search engines has been carried out (Spink et al., 1999; Spink et al., 2000; Spink et al., 2001; Jansen et al.

2000). These studies have investigated a number of variables related to search terms used by the public, including the number and types of search terms, the number of terms per query, search term subject categorisation, search term frequencies and co-occurrences, as well as search strategies and tactics. The results reveal that most people use few search terms, view few Web pages, and rarely employ advanced search features such as Boolean operators or relevance feedback.

Hsieh-Yee (2001) provides a thorough review of studies that have investigated web search behaviour of both children and adults in a wide variety of environments. She pointed out that researchers of web search behaviour have drawn upon earlier research into online search behaviour and have studied such variables as search terms, search reformulation, search patterns, tactical moves, search time and types of search task.

Summary of Search Term Selection Research

Table 2.1 illustrates key characteristics of the major search term selection studies discussed above. The studies reviewed have involved a variety of research populations. Nine studies investigated the search behaviour of professional searchers, two of which compared professional searchers with novice users. Three investigations focused on the search term selection patterns of end-users. A wide range of databases from different disciplines was used in these studies. Six studies used databases from the broad area of health and medical science, while the other six exploited social science, humanities, and library and information science databases for their investigations. The variables identified within the search term selection studies can be classified as follows: *searcher characteristics*: search experience, education, subject knowledge, and satisfaction; *search characteristics*: including search strategy, search tactics and search commands; *search term characteristics*: including free

text terms, thesaurus terms; and *search results characteristics*: such as relevance and effectiveness.

Research adopting a user-centred approach to search term selection demonstrates the complexity and importance of human decision making in the IR process. There is a growing interest in carrying out research into cognitive and behavioural aspects of users' search term selection and into the factors and variables that affect the process. Searching behaviour and the search process have been studied to explore issues associated with different types of user-system interaction and to discover how variables such as the user's initial request, search experience, environment, and domain knowledge affect the term selection process.

Since the search process in general and search term selection in particular involve a range of cognitive and behavioural characteristics, most studies have employed a combination of data gathering techniques to provide sufficient qualitative as well as quantitative data. The qualitative techniques derive mainly from disciplines within the social sciences, human-computer interaction and psychology. Data gathering tools utilised in these studies have included questionnaires (pre-search, during-the-search and post-search), interviews (both audio and video taped), the think aloud technique, transaction logging, and observation.

Most of the studies reviewed above have considered in varying degrees of detail, the role and influence of thesauri as one source for term selection. While some regarded thesauri as marginal or peripheral tools, others considered them to be substantial sources of terms which users can take advantage of during their online searching. Format and presentation of thesauri appear to affect their use. Most of the thesauri used by searchers and users in the studies noted were in printed format

and were thus considered to be external sources. Even in those studies where thesauri were available in electronic format, no specific attempt was made to examine and evaluate the user-thesaurus interaction as part of the term selection process.

Most of the studies reviewed focus on the term selection behaviour of professional searchers. Those studies specifically dealing with end-users give little attention to the role of thesauri as aids to search term selection. Rather their results are limited to suggestions on methods to facilitate the use of thesauri by end-users during the online search process. For instance Fenichel (1981) suggests the need for facilities such as hierarchical display of descriptors or "exploding" options to include all narrower terms of a descriptor, while Hsieh-Yee (1993) proposes that front-end software could incorporate features that encourage searchers to actively survey alternative search terms. Vakkari (2000) points to the need for system-provided synonyms and narrower terms as a means of facilitating search term selection and query reformulation.

As the number of electronic thesauri attached to IR systems has grown several interface features and facilities have been developed to aid users in the selection of search terms. However, very little research has been carried out into how end-users interact with these types of interfaces or into the ways in which these integrated thesauri affect search term selection (Jones et al., 1995; Beaulieu, 1997; Blocks et al., 2002). Research is required to shed light on the search term selection behaviour of various end-user communities who make use of IR systems that provide thesaurus and other terminological support for improving search performance. This research should not only address thesauri as term sources but also evaluate the impact of different types of user interface used to provide thesaurus-aided search facilities.

Author	Research population	Subject domain	Data collection techniques	Variables	Sources of search terms
Fenichel (1981)	Professional searchers & LIS students	Education	Questionnaire, search transcriptions	Search, searcher, search outcome, environmental	Thesaurus terms & free text terms
Fidel (1984)	Professional searchers	Health sciences	Observation, interview, think aloud, search protocol	Exploratory study – no specific variables reported	Thesaurus terms & free text terms
Fidel (1986)	Professional searchers	Life sciences	Observation, interview, think aloud, search protocol	Exploratory study – no specific variables reported	Thesaurus terms & free text terms
Saracevic et al. (1988)	Professional searchers & end-users	Medicine, sciences, social sciences & humanities	Video-taped interview, questionnaire, search records	Users, questions, searchers, searches, retrieved items	Oral & written problem statement, thesaurus
Fidel (1991)	Professional searchers	Medicine, sciences, social sciences	Observation, interview, think aloud, search protocol	Text-word ratio, thesaurus-neglect ratio, search keys, moves, subject area, environment	Thesaurus terms & free text terms
Bates et al. (1993)	End-users	The humanities	Transaction logs, interview	Database, search features, commands, terms and vocabulary	Free text statements and term relevance feedback
Hsieh-Yee (1993)	Experienced & novice searchers	Education	Transaction logs, think aloud, observe	Search experience, subject knowledge, search tactics	Thesaurus terms & free text terms
Iivonen (1995)	Professional searchers	Social sciences	Interview	Education, experience, environment, search term selection and strategy	Thesaurus terms & free text terms
Spink & Saracevic (1997)	Professional searchers & end-users	Medicine, social science, physical sciences & humanities	Transaction logs, video-taped interview	Search terminology, user characteristics, user satisfaction, search process and outcome	Questions, user interaction, searcher, thesaurus, term relevance feedback
Iivonen & Sonnen-wald (1998)	Professional searchers	Social sciences	Interview	Exploratory study – no specific variables reported	Thesaurus, domain & docs, indexing practice, search request, database, search experience
Sutcliffe et.al. (2000)	End-users	Medicine	Audio-video recording, think aloud	Query terms, query syntax, search strategies and effectiveness	Thesaurus terms & free text terms
Vakkari (2001)	End-users	Library & information science	Transaction logs, think aloud, interview, questionnaire	Knowledge of the topic, number and types of search terms, operators and tactics used	Free text

Table 2.1 Major studies which have adopted a user-based approach to thesaurus assisted search term selection

2.2.5 Query Expansion

Query expansion is defined as a stage of the IR process during which a user's initial query statement is enhanced by additional search terms in order to improve retrieval performance. As studies addressed specific stages of the search process the application of thesauri in this expansion and reformulation became an area of increasing interest. Query expansion is rationalised by the fact that initial query formulation does not always reflect the exact information need and request of the user; therefore, the query may often be enhanced by the addition of search terms in a manner that results in improved information retrieval.

Three types of query expansion are discussed in the literature, namely: manual, automatic and interactive (also known as semi-automatic, user-mediated or user-assisted). These approaches use different sources of search terms and a variety of expansion techniques. Beaulieu and Robertson (1996) have argued that the distinction between manual and interactive methods of query formulation is problematic, since both involve human intervention. The difference is that the manual approach does not include any consultation of the collection, while in the interactive approach the query is modified through a feedback process. In both cases, however, assistance can be sought from other sources, including a dictionary or thesaurus.

Spink (1994b) has commented that extensive research has been carried out on automatic and semi-automatic query expansion techniques. Automatic query expansion techniques exploit the text of a user's question and/or retrieved documents found to be relevant by the user, as input for techniques to derive a set of search terms to retrieve additional relevant documents. In interactive query expansion users are responsible for selecting from candidate search terms suggested by the retrieval

system. Several studies have been conducted on interactive query expansion to evaluate the ranking algorithms based on users' relevance judgement of candidate search terms (Efthimiadis, 1993; Efthimiadis, 1995), to study user interaction through graphical user interfaces (Beaulieu et al., 1995; Beaulieu, 1997), and to investigate the simulated users' term selection within the context of interactive query expansion (Harman, 1988; Magennis and Rijsbergen, 1997).

The human approach to query expansion research, "investigates the user's representation of their question and whatever tools e.g. thesaurus or experiences they use to extract or modify a set of search terms during query expansion" (Spink, 1994a). This approach stresses the importance of decision-making as well as behavioural and cognitive characteristics of users in reformulating and expanding their search statements.

Query Expansion Using Thesauri

Several studies have reported the construction and use of different types of thesauri as aids to the query expansion process. In general thesauri within IR systems can be categorised as belonging to one of three main types: standard manually constructed thesauri, searching thesauri, and automatically constructed thesauri.

Standard thesauri with hierarchical, equivalence and associative relationships have been widely used for search term selection and query expansion purposes. Much of the research in this area has focused on comparing the performance and effectiveness of controlled vocabularies versus free text terms in IR (Markey et al., 1980; Perez, 1982; Svenonius, 1986, Dubois, 1987; Cousins, 1992; Rowley, 1994; Muddamalle, 1998). These types of thesauri have also been incorporated as knowledge bases or interface components in several prototype expert

and intelligent systems to assist users in the process of search terms selection and query expansion (Efthimiadis, 1996).

Searching thesauri, also referred to as end-user thesauri, are defined as a category of thesauri enhanced with a large number of entry terms that are synonyms, quasi-synonyms or term variants which assist end-users to find alternative terms to add to their search queries (Perez, 1982; Piternick, 1984; Bates, 1986; Cochrane, 1992). A number of searching thesauri have been designed and developed (Anderson and Rowley, 1991; Lopez-Huertas, 1997; Knapp et al., 1998; Nielsen, 2001) and have been evaluated in query expansion research (Kristensen and Jarvelin, 1990; Kekäläinen and Jarvelin, 1998; Kristensen, 1993).

The design and testing of several types of automatically constructed thesauri has also been extensively reported in the literature. A number of researchers have constructed co-occurrence based thesauri to evaluate the performance of thesaurus-based query expansion (Qiu and Frei, 1993; Schutze and Pedersen, 1997). Using a laboratory environment and the TREC test collections, these studies resulted in a slight improvement in retrieval performance. General-purpose thesauri such as WordNet have also been evaluated in the query expansion process but have demonstrated little difference in retrieval effectiveness (Voorhess, 1994). Thesauri constructed automatically using a linguistic approach have also demonstrated a marginal improvement in retrieval performance (Jing and Croft, 1994). Combining different types of thesauri for query expansion has shown better retrieval results than using only one type of thesaurus (Mandala et al., 1999).

Automatically constructed thesauri have also been evaluated in user-oriented environments (Chen and Dhar, 1991; Chen et al., 1995; Chen

and Ng, 1997). In addition, some researchers have found that the integration of automatically and manually constructed thesauri has a positive effect on the query expansion process (Chen and Ng, 1995; Schatz et al., 1996; Chen et al., 1998; Ding et al., 2000).

The research reviewed in the following section focuses on those studies that have investigated thesaurus-based query expansion using standard domain-specific thesauri in a user-oriented environment.

2.2.6 Thesaurus-aided Query Expansion: User-centred Studies

Research on user-assisted query expansion using domain-specific thesauri can be divided into three categories depending on the extent to which users are involved in the search and retrieval process:

- users involved only in providing requests and relevance judgements;
- user involvement mediated by professional searchers; and
- users operating in an interactive environment.

The following sections review studies conducted in each of these categories.

Users Involved Only in Providing Requests and Relevance Judgements

In the studies reported in this section, users were not involved in the actual search process; searches were typically performed by a professional searcher who also carried out any thesaurus-aided query expansion. Kristensen and Jarvelin (1990) studied the effectiveness of a small searching thesaurus, with 328 terms, on recall and precision in a full-text database. The test environment was an operational database with around 34,000 newspaper articles relating to economic issues and

the queries were elicited from five journalists. The researchers performed the searches while relevance judgements were undertaken by the journalists. Each query was searched in three distinct modes: basic search, synonym search, and related term search. The basic search included only the journalists' initial query statements. Expanding original queries with synonyms using a searching thesaurus significantly increased the number of new relevant records. Query expansion using related terms also increased the recall rate but led to a marked decline in precision. Thus related terms can be used as query expansion terms if high recall is required. They concluded that a searching thesaurus could improve results in free-text searching of a full-text database.

Following this study, Kristensen (1993) investigated the effects of the searching thesaurus on recall and precision in a full-text database using an expanded set of five distinct search modes: basic search, synonym search, narrower term search, related term search, and union of all these searches. In this study, an operational database of 227,000 newspaper articles, a test thesaurus of 1573 terms, and 30 queries elicited from journalists constituted the test environment. The researcher carried out all modes of search, with relevance judgement again being supplied by the journalists. The study concluded that the effect of a search-aid thesaurus was substantial and improved retrieval with twice as many relevant documents being found using the union search as compared to the basic search, with only a 10 percent decrease in precision. Each of the three expanded search modes retrieved several unique articles, but it was found that the expansions using synonyms and related terms performed better in terms of relevance than those using the narrower terms. The findings showed that a searching thesaurus was clearly a recall-enhancing tool and the author suggests that the active involvement of end-users in the term selection process could enhance the levels of precision.

User Involvement Mediated by Professional Searchers

To evaluate thesaurus-aided query expansion in a more user-oriented setting, some researchers have studied the selection of search terms for query expansion by users in mediated search environments. Spink (1994a) for instance examined the selection and effectiveness of search term sources while observing used-based query expansion to provide guidelines for IR system enhancement. The users' written question statements and pre-online interviews, as well as the search logs created during online interaction between 40 users and professional searchers, were collected. The study identified five different sources of term selection for query expansion: user question statement, user interaction (suggested by user during interaction with intermediary or online search), intermediary, thesaurus, and term relevance feedback (TRF) based on terms extracted by users after examination of retrieved documents. Spink also evaluated the effectiveness of these sources in retrieving relevant items with a special focus on term relevance feedback. The thesaurus was rated as the third most effective source of search term after the user's question statement and user interaction with the intermediary. The results showed that the TRF search terms were more effective in retrieving relevant items than search terms suggested by the professional searchers or those selected from the thesaurus. TRF terms that proved effective in retrieving relevant records were largely selected from the title and descriptor fields of the records viewed. This finding suggests that descriptors, which will normally be terms in the thesaurus, contributed significantly to the retrieval of relevant records. Thus, due to the fact that end-users did not have any direct interaction with the thesaurus as a main source of term selection, the value of thesaurus-based terms may have been underestimated. The study concluded that users' written question statements, terms derived during the interaction between users and professional searchers, and terms selected from the title and

descriptor fields of the retrieved records were the most effective sources of search terms for query expansion.

Greenberg (2001a, 2001b) investigated the effectiveness of different thesaural relationships for automatic and interactive query expansion in an operational environment. She aimed to explore how end-user search terms can be mapped to a thesaurus and, if there are additional thesaurus terms representing some type of semantic relationship, to investigate which are good candidate terms for interactive and automatic query expansion. The test environment consisted of the *Proquest* controlled vocabulary, ABI/Inform (a business periodical database), and forty-two queries obtained from business administration students. The study collected data from user profile questionnaires, relevance judgements, and a post-evaluation questionnaire. In evaluating automatic query expansion using thesauri, different query treatments were reported in the study. These 'treatments' differed in terms of the sources used to expand the query. She included the participants' initial search statement and five treatments involving the thesaurus: mapped terms, narrower, broader, related, and synonymous terms. In the interactive query expansion process, the participants were presented with lists of thesaurus terms related to their initial queries and were asked to select search terms that they thought would have been useful. In both the automatic and interactive cases the researcher developed the search strategy and performed all the searches while end-users were mainly responsible for providing real information requests together with relevance judgement. The findings suggested that synonyms and narrower terms yielded more precise results and are good candidates for automatic query expansion because they increase recall with a minimum loss in precision. All semantic relationships increase recall if applied in automatic query expansion. If precision is required, narrower terms can be good candidates for interactive query expansion. Related terms can

be better candidate terms for interactive query expansion than for automatic query expansion. The results also demonstrated that end-users can have a significant impact on the precision of the results when selecting terms via interactive query expansion. This study indicated that semantic relationships in standard thesauri can have an impact on retrieval performance and can be used as a source for query expansion.

Users Operating in an Interactive Environment

The research reviewed in the previous two sections covers studies with only partial user involvement in thesaurus-aided search term selection for query expansion. There are a few studies that have investigated such selection behaviour by end-users in an interactive environment. Jones et al. (1995) for instance, investigated user interaction behaviour with a thesaurus as a source of query expansion terms to identify strong patterns for possible rule- or weight-based systems for term expansion and to compare the effectiveness of queries enhanced by thesaurus use with that of the original query terms. The INSPEC thesaurus, 39 users with real information needs and OKAPI, a ranked output IR system, were the experimental environment within which the users' thesaurus navigation and term selection behaviour was recorded. The interface was implemented in such a way that following the entry of the original query users were shown exact or partially matched thesaurus terms and encouraged to browse and select thesaurus terms for expansion. The researchers examined the number and types of terms seen and selected by users. The analysis showed that the majority of terms retrieved by thesaurus navigation came through the association relationship, perhaps reflecting the fact that INSPEC had more associative links than hierarchical ones. The focus of the study was on the performance of thesaurus terms in relevant retrieval. A comparison of the results generated by original queries with those generated by the selected thesaurus terms was made based on the users' relevance judgements

(using only author, title, and journal) and showed no significant differences in the retrieval performance. The study also evaluated the retrieval performance of hybrid searches, which included both the original and controlled terms. The hybrid search retrieved more records than the original search although it did not show a marked improvement in overall performance. The researchers concluded that thesaurus-based query expansion may increase recall and uncover additional relevant documents, but will not improve queries which are already quite fully specified. It was also suggested that users obtain good results if they have a large number of terms from which to select, and that since their choice of terms will differ it is infeasible to design a generic automatic expansion procedure. Rather tools should be provided to aid the end-user in thesaurus navigation. User feedback also indicated that thesaurus navigation was a useful and informative activity. It was also found that depth and coverage of the thesaurus and the quality of the user interface were important issues in determining the degree of enhancement to the query expansion process.

Beaulieu (1997) reported on experiments conducted to evaluate different interfaces which support query expansion, again based on the Okapi search engine. The experiments were carried out using an online library catalogue and also the INSPEC database and thesaurus. Observation, talk-aloud, online pre-search and post-search questionnaires were used together with transaction logs to gather data. Three types of interfaces supporting automatic and interactive query expansion were tested by examining user interaction behaviour. Thesauri were used in different ways within the interface design, as tools for supporting interactive query expansion. The results showed that both the explicit and implicit use of a thesaurus (using interactive or automatic query expansion respectively) can be beneficial. It was found that the overall number and specific presentation of candidate terms for query expansion was important. For

instance, the way that thesaurus terms and terms extracted from the documents were displayed affected the ultimate set of search terms selected by the user. The study also suggested that the different cognitive styles adopted by users when seeking information were an important issue in designing interfaces that could effectively support query expansion.

Summary of Thesaurus-aided Query Expansion Research

Table 2.2 provides an overview of some key characteristics of the query expansion studies discussed above. Two of the studies adopted a laboratory experiment approach, where end-users were only involved in providing search requests and relevance judgements (Kristensen and Jarvelin, 1990; Kristensen, 1993). In two other query expansion investigations, end-users' search behaviour was examined in an operational setting (Beaulieu, 1997; Jones et al., 1995). The remaining pair of query expansion studies considered a mediated environment where professional searchers interacted with end-users (Spink, 1994a; Greenberg, 2001b). Within these studies, three investigations utilised economic and business databases, two involved computer and information science databases, while the final study covered a range of subject domains.

The results of these studies demonstrate the usefulness of thesauri both in terms of providing users with alternative search terms for query expansion and in improved retrieval performance. However, these studies focused primarily on search outcomes rather than the search process. Interactive query expansion research (Beaulieu, 1997) has shown that user interface design can play a major role in encouraging the use of thesauri as query expansion term sources. This has implications for the design and evaluation of interfaces enhanced with thesauri to support

query expansion. Efthimiadis (2000) has suggested that during query expansion, a thesaurus could be used to display the relationships of the selected terms to other terms. This could be achieved, for example, by displaying the hierarchical tree to which a term belongs (as in the INSPEC or MESH tree displays) or by presenting broader, narrower or related terms on screen for users to browse and make selections from. As end-user searching becomes more prevalent, research on user - thesaurus interaction within the context of the information searching process is also needed to evaluate the ways in which various end-user communities use and interact with thesauri and how this affects their term selection for query expansion. Such research must take into account the user's attitudes, cognitive aspects of the search process, and the mechanisms by which the user may select terms, issues little investigated in the reviewed literature. Research also suggests that the coverage and richness of thesauri play a significant role in their contribution to the users' term selection for query expansion (Fidel, 1992; Jones et al., 1995). Given the fact that few domain-specific thesauri have been evaluated in terms of their coverage and performance for query expansion, research needs to be carried out to evaluate thesaurus-aided query expansion in a range of subject domains.

Author(s)	Research population	Subject domain	Database	Thesaurus
Kristensen & Jarvelin (1990)	End-users	Economics	Finnish newspaper articles	In-house search-aid thesaurus
Kristensen (1993)	End-users	Economics	Newspaper articles	In-house search-aid thesaurus
Spink (1994a)	Professional searchers and end-users	Various subjects	Dialog bibliographic	Not specified
Jones et al. (1995)	End-users	Computer, electronics, & information science	INSPEC	INSPEC
Beaulieu (1997)	End-users	Computer, electronics, & information science	INSPEC	INSPEC
Greenberg (2001)	End-users	Business	ABI/Inform	ProQuest controlled vocabulary

Table 2.2 Key studies which have considered the role of users in thesaurus-based query expansion

2.2.7 Conclusion

This section of the literature review has provided a survey of user-centred studies on search term selection and query expansion. In particular the use of domain-specific thesauri as sources of term selection has been examined. The studies reviewed were classified according to the type of user involved and it was seen that most of the search term selection studies investigated the behaviour of professional searchers. This was due to the fact that a number of the studies were carried out at a time when mediated online searching was common practice. In addition, most IR systems did not have integrated online thesauri and so professional searchers were the main users of these tools. Research demonstrated

that user characteristics, including cognitive and behavioural factors, are important in analysing end-user search term selection.

Query expansion studies by contrast have been more concerned with the end-user, although in most cases end-users were not involved in the actual search process. Moreover the focus of thesaurus-based query expansion research has been on retrieval performance rather than the search process and its cognitive aspects. The results of these studies demonstrated that thesauri have the potential for improving retrieval performance and can be used for both automatic and interactive query expansion. In terms of subject focus the review has identified a gap and has indicated that in both the area of search term selection and query expansion there is a need for studies which consider users and databases from disciplines within the basic and life sciences.

While there exist a number of studies exploring the general information seeking behaviour of agricultural scientists and veterinarians (Nweke, 1995; Majid et al., 1999, 2000) none of them have investigated veterinarians' information retrieval behaviour while interacting with an IR system. The present study has therefore targeted this particular end-user community in order to evaluate their behaviour as it relates to search term selection and query expansion using thesauri.

Two major trends in modern information retrieval justify additional research into thesaurus-assisted search term selection and query expansion by end-users. First, end-user searching is widely becoming universal owing to the developments associated with the Internet and in particular the World Wide Web. To provide better systems, research that explores the behaviour and attitudes of the different types of end-users who access thesaurus-enhanced IR systems is needed. User behaviour incorporates both cognitive and physical aspects such as the way in

which users relate thesaurus terms to their search topic and the use of specific interface features respectively. The relevant attitudes of end-users in this context would include such subjective issues as their satisfaction with a result set and its relevance or their motivation to utilise facilities provided by the thesaurus. Second, a growing number of commercial retrieval systems have recently incorporated thesauri into their search interfaces in an attempt to encourage users to enhance their queries. Shiri et al. (2002b) have reviewed six commercial web-based IR systems that provide thesaurus-enhanced search interfaces.

This development opens avenues for more research into the usefulness and usability of these systems and the extent to which they do in fact assist end-users in selecting search terms and in expanding queries. Studying end-users' interaction with online thesauri will extend our knowledge of user behaviour and provide guidance for the design and implementation of better systems.

2.3 Thesaurus-enhanced Search Interfaces

This section of the literature review surveys the literature covering efforts to integrate standard thesauri as part of IR search interfaces with the aim of assisting users in selecting search terms for query formulation and expansion. It provides details of features and facilities included in thesaurus-enhanced search interfaces. The emphasis of this review is on those interfaces which have integrated standard thesauri with equivalence, hierarchical and associative relationships, while those providing links to automatically or statistically constructed thesauri are not considered. Since the focus of this review is on search and retrieval interfaces enhanced by thesauri, it does not discuss issues related to thesaurus management software.

The first major part of this section deals with the rationale behind integrating thesauri as part of search interfaces. This looks at a number of studies on searching behaviour, IR interface evaluation, search term selection and query expansion research providing evidence for the incorporation of thesauri into interfaces. The second part reviews research-based prototypes as well as commercial web-based interfaces enhanced with thesauri. It highlights thesaurus-enhanced interface features and the way in which they can help users during the search process. A list of thesaurus-related features, which can be implemented in the search interfaces, is provided at the end of the section.

2.3.1 Introduction

The selection of search terms for query formulation and expansion is one of the key issues in the IR process. The subject approach in general and the use of thesauri in particular have been recognised as a method of enhancing the search term selection process. Traditionally, thesauri were used by search specialists for selecting alternative search terms to improve results. Recent developments in end-user searching and the

wider availability of online IR systems together with advances in user-centred interface design have opened new horizons for utilising thesauri as search aids for end-users.

The significance of the interface in supporting information seeking in general and search term selection in particular has been emphasised in IR interaction models (Belkin and Cool, 1993b; Ingwersen, 1996; Saracevic, 1997). At the heart of all of these models lies the process of query formulation. As emphasised by Saracevic (1997) selection of search terms for query formulation is a dynamic, interactive process that calls for a wide variety of interface features and facilities to support users.

One of the issues in term selection is how users select terms for query building and how these terms match to the system vocabulary. A number of researchers have elaborated on the variability of search terms provided by different users while interacting with IR systems. Furnas et al. (1987) for instance, have studied vocabulary problems in using computer applications and cited the large variability of word selection by different users as evidence for the need to develop systems which recognise a rich variety of vocabularies. They also concluded that there were many avenues for interface improvements to incorporate vocabulary components. Bates (1986,1998) has discussed the improvement of both interfaces and vocabularies of IR systems to support users in the selection of search terms. Her well-cited model of a vocabulary-enriched interface, which she termed the Front-end System Mind (FSM), provides end-users with a wide range of search terms, alternative term displays and various approaches to term selection. She refers to the concept of an end-user thesaurus in order to highlight the importance of providing users with a large number of alternative terms. She elaborates on this concept by noting that users will be able to perform powerful searches if an initial

term or topic submitted results in a screen full of term possibilities, related subjects, or classifications from which the user can make a selection.

Milstead (1997) stated that thesauri are underused by searchers due to the fact that database providers do not make them readily available. Some are available only in paper form while those provided online often have little or no user support in the interface. She suggested that the provision of effective thesaurus navigation tools within the interface will improve searcher access and use.

The importance of thesaurus support for effective searching has also been stressed by researchers in the field of human-computer interaction. Shneiderman (1998) for instance, in discussing user interface features for searching textual databases, puts forward a four-phase framework: formulation, initiation of action, review of results, and refinement. He suggests that in the search formulation phase, interfaces should provide broader, narrower and synonymous terms from a thesaurus to help users clarify their search.

2.3.2 Earlier Reviews of Information Seeking Interfaces

User interfaces to IR systems that support information seeking processes have been widely discussed in the literature. Marchionini (1992) for instance, provides a description of essential features of interfaces to support end-user information seeking and suggests five information seeking functions, namely problem definition, source selection, problem articulation, result examination, and information extraction. He argues that much of the interface work has focused on problem articulation (including query formulation) and that other functions need to be investigated in designing information seeking interfaces. Marchionini and Komlodi (1998) discuss the evolution of interfaces and trace research and

development in three areas namely information seeking, interface design and computer technology. They provide a brief review of interfaces to online IR systems as well as Online Public Access Catalogues (OPACs) and discuss the new generation of user interfaces influenced by the emergence of the World Wide Web. They conclude that interface design has become more user-centred and the trend is toward more mature interfaces that support a range of information seeking strategies. Savage-Knepshield and Belkin (1999) have characterised trends related to interface design challenges within the context of IR interaction over the last three decades. They suggest three major eras – which they refer to simply as the early years, the middle years, and the later years – to provide a description of the types of interfaces designed in these periods. Command language interfaces provided the main approach in the early years. In the middle years menu-driven and form fill-in interfaces, which were more appropriate for novices and casual searchers, became the dominant interface type. In the later years, users and their information needs became the focus of the most complex interface design challenges. This period is characterised by use of the natural language and direct manipulation user interfaces. They note that the degree of interaction between the searcher and the IR system has dramatically increased and much research is still required to meet the challenges in interface design for IR interaction.

Looking at both human-computer and IR issues, Hearst (1999) discusses user interface support for the information seeking process and describes the features of these interfaces that aid such processes as query formulation and specification, viewing results and interactive relevance feedback. She describes a number of graphical user interfaces that provide information seekers with a wide range of approaches to specify, view, analyse and evaluate both query and document within the context of IR systems. Interfaces that support the formulation of both Boolean

and natural language queries, as well as those providing categorical and subject support, are examples of the reviewed interfaces. Her review provides a few examples of thesaurus-enhanced search interfaces whose details will be discussed in the section on research-based interfaces. She concludes that there is an increasing interest in taking the behaviours of individuals into account when designing interfaces. In this rapidly developing area she suggests that improvements in the interface will lead to better search results and more facilities for information creators and users.

2.3.3 The Rationale for the Integration of Thesauri into Search Interfaces

Several empirical investigations into searching behaviours, IR interfaces, and term selection and query expansion have stressed the necessity of some form of thesaurus support at the interface level to aid users in selecting search terms. These are discussed in the following sections.

Searching Behaviour Studies

In studying information seeking strategies of novice users searching a full-text encyclopaedia, Marchionini (1989) found that users had difficulties in selecting terms for query formulation and concluded that the addition of a thesaurus or usage-sensitive search aid would provide for more efficient and effective searches. Hsieh-Yee (1993) investigated the effect of search experience and subject knowledge on the search tactics of novice and experienced searchers and pointed out that they used a limited number of sources for term selection. She also suggested that front-end software should include features that encourage searchers to view possible search terms and take full advantage of a thesaurus. She proposed the incorporation of some interface features which would support users in search term selection, such as the "see" references to

translate the user terms into terms acceptable to the system, automatic prompting for the inclusion of synonyms and access to a large online thesaurus. Another empirical study undertaken by Sutcliffe et al. (2000) investigated end-user searching behaviour on the MEDLINE database, part of which focused on query construction and term use. They reported that searchers had difficulty in choice of terms even though MEDLINE has thesaurus and term suggestion facilities. They argued that educating users about these facilities only goes part of the way to solving the problem. They noted that since term suggestion and the use of thesaurus facilities required several steps, many searchers did not make the effort required to use them effectively. The researchers suggested that there was a need for a more *active* thesaurus facility to support query articulation and that such a facility should explicitly incorporate guidance for thesaurus usage as part of the query dialogue. They also pointed out some of the usability problems relating to the thesaurus and term suggestion facility.

IR Interface Evaluation

Another line of investigation, providing evidence for incorporating thesauri into interfaces for supporting search term selection, deals with IR interfaces. Bates (1990b) proposed a number of search term tactics and strategies, which can be implemented in the interfaces to IR systems. These specific term tactics include SUPER for broader search terms, SUB for narrower search terms and RELATE for related terms which are comparable to the conceptual structure of thesauri and can provide the user with alternative approaches to search term selection.

Borgman et al. (1989) studied the information needs and information seeking habits of researchers in the energy domain and designed and evaluated an automated assistance program to act as a front-end to aid end-users in searching energy databases. The study showed that there

was a need to provide search term support for end-users as they had considerable difficulties in selecting search terms. They concluded that some of the vocabulary problems could be attributed to the lack of online access to the energy database subject thesaurus to find out the alternative broader, narrower, and related terms. However, they also commented that a significant proportion of the vocabulary problems could be resolved through having index browsing capabilities. There was no indication as to whether having an index browsing facility in place would eliminate the need for thesauri or not.

The behaviour of two types of users with two types of IR interfaces was studied by Meadow et al. (1995) to investigate the differences in performance and attitudes of the users. Professional librarian searchers and graduate library students constituted the research populations and search term selection was one of the major factors controlled and analysed through the study. The two interfaces considered in the study were Dialog with a procedural language and the OAK menu-based interface to Dialog. In the case of both types of interface, the process of editing/refining queries and term selection was perceived to be a problem by both groups and they felt the need for vocabulary assistance. Search specialists as well as students required more help in the conceptual aspects of the search process. Responses to questions relating to the search term selection process indicated that both groups would have liked to have had access to a thesaurus or dictionary.

Search Term Selection and Query Expansion Research

Investigations into search term selection and query expansion also have implications for the role of thesauri in the search process and their accessibility within the interface. Fidel (1991, 1992), in a series of experiments on the search term selection behaviour of professional

searchers, found thesauri to be major tools which assisted them in various stages of selecting search keys and improving search results. She concluded that database producers and search-system vendors should provide easy, flexible, and inexpensive online access to thesauri. Among her suggestions for the mechanisms supporting search term selection using thesauri were the provision of expanded lists of synonyms, automatic execution of inclusive searches (using an *exploding* or *cascading* approach), and the display of hierarchical relationships for a descriptor at any point during the search process. She also proposed the concept of an *intelligent* interface through which the vocabulary of a searcher could be matched to the vocabulary of a thesaurus.

Vakkari (2000) investigated the changes in search terms and tactics of eleven students during accomplishing a research proposal for master's theses. The results demonstrated that as students' understanding of the task is growing, they tend to select narrower and synonymous terms. He suggested that the provision of terminological support would enable them to differentiate the topic, to choose narrower terms and to enhance search results.

Query expansion research has also shown that the integration of thesauri into interfaces provides support for users in choosing search terms during the search process. Beaulieu (1997), Beaulieu et al. (1997) and Jones et al. (1995) in a series of experiments on designing various interfaces to the Okapi search engine for automatic and interactive query expansion have demonstrated that both explicit and implicit use of a thesaurus during automatic and interactive query expansion can be beneficial. They also stated that while useful terms can be selected through automatic expansion, terms explicitly selected by users were also valuable. Many of

the users in the experiment on thesaurus-based query expansion found the process of thesaurus navigation in itself to be useful and informative.

In a user-centred investigation of interactive query expansion, Efthimiadis (2000) found that most of the query expansion terms selected by users were identified as being hierarchically related to the initial query terms. He commented that this finding had implications for designing user interfaces in a manner such that a thesaurus could be used to display the hierarchical tree to which the term belonged and provide broader, narrower or related terms from which the user could browse and select.

In addition to the above research, there are a number of studies undertaken within the context of the TREC (Text REtrieval Conference) interactive track which focus on search term suggestion facilities and related interface issues. Belkin et al. (2001) report a series of TREC investigations on interactive query reformulation and the role of term suggestion facilities within the IR interface. While these studies do not refer to thesauri as search term sources, their results show that user-controlled term selection is more effective than automatic query expansion and that users are able to understand and learn how to use search term suggestion facilities provided by IR systems.

Joho et al. (2002) have examined the effectiveness of a hierarchical presentation of candidate terms for query expansion in comparison to the conventional format of an ordered list. They found that users who expanded their queries using the hierarchical list spent less time and chose fewer terms to complete the query expansion task than those who used the list format. The research also concluded that different presentations of expansion terms affect the expansion task.

2.3.4 Thesaurus-aided Query Formulation and Expansion: Early Systems

To provide some background to the evolution of thesaurus-aided search systems, a brief description of earlier attempts to exploit thesauri for supporting query formulation and expansion is presented. The growing application of online or electronic versions of domain-specific thesauri for query formulation and expansion can be traced back to the late Seventies when a number of IR researchers began to develop prototype systems in order to explore ways of aiding user search within IR systems. The development of expert system and artificial intelligence technologies in 1980s provided the grounds for a growing interest in applying thesauri as the knowledge bases of a number of expert systems and intelligent front-ends. Efthimiadis (1996) provides a detailed review of these thesaurus-enhanced systems, most of them using expert system techniques, which were designed and developed to assist users in formulating and expanding queries in one way or another. Many of these systems embedded thesauri as part of their search facilities, which provided users with a number of alternative search terms from which to choose. Some of these systems used mapping techniques for matching user submitted terms with their thesaurus knowledge base and displayed hierarchical structures associated with the entered term. Most of these expert and intermediary systems used standard thesauri such as MeSH and INSPEC to provide either thesaurus browsing or thesaurus mapping capabilities. The section that follows builds on this review by looking at recent developments in thesauri-enhanced systems and their interface features.

2.3.5 Review of Thesaurus-enhanced Search Interfaces

This section reviews research on IR interfaces that have integrated domain specific standard thesauri as browsing or searching facilities in either prototype or operational IR systems. It should be noted that

thesaurus building and editing software, which is widely used by indexers for thesaurus construction and maintenance, is not discussed. This review focuses on interfaces that make use of thesauri as a means of improving users' performance in term selection at the search and retrieval stage.

When adopting thesauri for incorporation into an interface two main types of thesauri must be considered. One type, typically referred to as *standard* thesauri are usually used as controlled vocabularies for indexing and retrieval. Most of the research-based prototypes reviewed here have adopted this type of thesauri within their search interfaces. The other major type, known as *search* thesauri or *end-user* thesauri, are usually enhanced with a large number of equivalent terms such as synonyms, quasi-synonyms and term variants. These thesauri function as tools which assist end-users in finding alternative terms to add to their search queries and thus enhance retrieval performance. This latter category includes the interfaces proposed by Bates (1990a) and Schatz et al. (1996).

There is substantial documentation on designed, prototyped and operational thesaurus-enhanced interfaces that have adopted different approaches in thesaurus incorporation. Two general types of interface are recognisable within the related literature, those designed for research-oriented and commercial systems, and this division is used in the sections which follow. The latter category were originally delivered using the medium of CD-ROM but most have recently been made available on the Internet, and it is the Web-based versions of the commercial systems which are considered in Section 2.3.5.2.

Research-based Interfaces

This section examines a number of research-based prototype interfaces which address a number of academic issues involved in the incorporation of thesauri into interfaces. A range of textual and graphical approaches have been taken to utilise thesaurus in the interface.

A great deal of work on the incorporation of thesauri into interfaces has been carried out over the past 15 years by Pollitt and his colleagues. In CANSEARCH and MenUSE two hierarchically linked menu-based interfaces, Pollitt et al. (1987, 1988, 1993, 1994a) adopted a novel approach by showing the thesaurus terms on the interface as the means by which users could select terms and formulate queries. The display of top terms in the thesaurus together with the number of terms under each main concept and, importantly, the number of the related references in each case were key features of both these systems. The AND operator was available to allow the user to carry out searches using a combination of thesaurus concepts. Through these interfaces users can progressively refine their queries from top terms of the thesaurus to the specific concepts under each concept.

Moving from a character-based interface to view-based searching (using interface features such as those made available in the Microsoft Windows and Apple Macintosh operating systems) Pollitt et al. (1994b, 1996a, 1996b, 1997) designed the HIBROWSE (High Resolution Interface for Database Specific Browsing and Searching) system as an extension to the MenUSE interface with more facilities for bibliographic databases. The system was implemented as a front end to a number of bibliographic databases including Medline, INSPEC, EPOQUE (European Parliament Online Query System) and EMABSE. While MenUSE provides sequential access to menus for query formulation, the HIGHBROWSE system

presents a multi-window search interface with the thesaurus at the centre. The interface presents views onto the database using navigable hierarchies of subject descriptors from thesauri and provides simultaneous access to different bibliographic fields. One of the notable features of the HIBROWSE interface is its ability to simultaneously show the thesaurus terms associated with two or more search facets in a single view. The interface caters for navigating up and down the hierarchy for further query refinement. While both MenUSE and HIBROWSE have extensively utilised thesaurus knowledge, the HIBROWSE system provides more dynamic and multidimensional access to the content of the databases.

To put into practice her ideas on thesaurus-enriched interfaces, Bates (1990a) proposed the design of a subject search interface and online thesaurus. The design features allow for mapping user entered terms to those of the thesaurus and the use of OR operator for combining all thesaurus terms selected by users. Displaying alternative thesaurus terms without any hierarchical order or term type definition is the main feature of the proposed interface. In using this type of interface, it is assumed that users do not need to have any idea of the thesaurus and its relationships and are only provided with a number of terms close to their own search terms. Pollard (1993) has reported the design of an experimental hypertext-based browsing interface for the ERIC thesaurus as a tool for subject access to the related bibliographic database. The user interface consists of two scrollable windows and a typical search session starts with a keyword search. The user then is guided to the details of the terms such as different thesaurus terms from which he can select terms for searching the database. The interface uses hypertext features for navigating through different parts of the thesaurus.

Hyperline (Agosti et al., 1991,1992) is a conceptually enhanced interface which makes use of the INSPEC thesaurus and hypertext features to provide concept browse and navigation facilities. The system, developed for the Information Retrieval Service of the European Space Agency, employs a mapping technique for matching words entered by users with thesaurus terms after which the interface shows candidate concepts for browsing. The thesaurus browsing consists of two stages. In the first stage, the interface shows user input and related thesaurus terms from which the user can make a selection. In the second stage the user is provided with the details of the thesaurus term selected including narrower, broader, related, top and *Used For* terms as well as the number of references indexed by that term. The system also provides a search term pool where users can save their browsed and navigated terms. Having access to the documents and thesaurus at all stages of navigation is a major feature of this interface, which caters for moving from thesaurus concepts to documents and vice versa.

In designing BRAQUE (BRowsing and QUEry formulation), an interface to support different information seeking strategies, Belkin et al. (1993a) employed Hyperline to provide a thesaurus browsing facility as one type of information seeking strategy. In their model of information seeking strategies, the use of thesauri or what they termed "meta-information" can be considered to be one such strategy. Using the Hyperline facility, BRAQUE supports various types of browsing and searching strategies which utilise thesaural knowledge.

A different approach was adopted by Johnson and Cochrane (1995) and Johnson (1997), who developed a prototype of a completely hypertextual interface to the INSPEC thesaurus. KWIC (Keyword-In-Context) and KWOC (Keyword Out of Context) provide alternative ways of accessing

the thesaurus to support users in navigating the thesaurus. The interface displays both hierarchical and alphabetical lists of terms and shows the current term around which its related terms are displayed. The interface uses + and - icons to display broader or narrower terms while at the same time taking advantage of visual features to display *Related Terms* to the term selected by the user in an adjoining area of a unified interface. All terms shown on the interface are clickable and can be selected for searching the database. Schatz et al. (1996) utilised this interface to design a prototype of a system which allows users to simultaneously combine terms offered by different suggestion techniques such as thesauri and co-occurrence list.

Shapiro and Yan (1996) also reported the design of an interactive thesaurus prototype for query building using hypertext features. The interface provides for term input as well as term browsing in a series of HTML pages. The different types of thesaurus terms can be selected using checkboxes. The interface can also display the full set of terms collected while the user has been browsing through the thesaurus.

In a project named CILKS (City Interactive Linguistic Knowledge System), Jones and Hancock-Beaulieu (1994) and Jones et al. (1995) designed a thesaurus-enhanced search interface to exploit thesaurus knowledge for query enhancement in a probabilistic IR system using knowledge-based techniques. The interface provides a number of facilities to ease the term selection process. The query entered by the user is completely or partially matched with the thesaurus terms and once a thesaurus term is chosen related terms are shown in separate narrower, broader and related sections. For each term, the number of matching terms, the number of documents indexed by that term and the number of *Related Terms* are shown on the interface. Users can have different search facets to

separately search in the thesaurus and selected thesaurus terms can be placed in a *pool* for further selection, deletion or expansion. The simultaneous display of all original and selected terms as well as the list of retrieved documents are key features of the interface. As some matching terms are *Lead-in*, the interface also shows the preferred terms for possible substitution. This type of interface has been specifically designed to meet the requirement of query enhancement and expansion.

Beaulieu (1997) has reported the design of an enhanced thesaurus-aided interface for supporting query expansion which provides simultaneous access to both thesaurus mapped terms as well as search results in a single window. The user is provided with a facility for adding the thesaurus mapped terms to the query or navigating the thesaurus. In addition to taking advantage of the features mentioned in the CILKS project, this interface also provides the user with access to the full text of each of the items in the retrieval set.

In addition to the largely text-based thesaurus interfaces noted above, a few prototype interfaces have utilised graphical as well as two or three dimensional category hierarchies using the MeSH Thesaurus. TraverseNet (McMath et al., 1989) is a graphical interface which allows for the graphical display of a hierarchical IR structure and its associated documents using a root node and children nodes attached. Queries can be formulated using thesaurus terms by selecting hierarchy terms and placing them in the query. The thesaurus is shown as a circle with children circles around which are depicted as semantically related terms. MeSHBrowse (Korn and Shneiderman, 1995), a prototype interface for browsing the MeSH thesaurus, employs a concept space approach and a node-link tree diagram of the concept space. Hidden inter-relationships of the terms are revealed once a node is clicked. The graphical nature of the

interface allows only for related categories of terms to be displayed in a two-dimensional tree structure. Cat-a-cone (Hearst and Karadi, 1997) utilises a three-dimensional graphical interface which shows all of the top level categories initially and allows for the user to control the subsequent expansion. An alternative mode of interaction is to have the user type in a category label and see which parts of the hierarchy match or partially match that label. The interface also caters for a kind of relevance feedback by suggesting additional categories. Users can jump easily from one category to another and can search on multiple categories simultaneously. Visual MeSH (Lin, 1999) is a graphical interface developed to interact with the MeSH thesaurus and Medline. It allows the user to look up MeSH terms in a click-and-choose environment and assists users in exploring the MeSH terms through providing several views of the concept including tree view, neighbour view and map view. On any of the views, the user can double-click on a term to select it.

Déjà vu (Gordon and Domeshek, 1998) is a thesaurus-enhanced search interface which takes advantage of the Library of Congress Thesaurus of Graphical Materials to provide a browsing facility for retrieval in a catalogue of digital media. The interface utilises a knowledge representation technique to group subject terms into clusters or *Expectation Package*, which capture the common-sense knowledge of library users. The interface consists of two main sections. The top section shows the thesaurus subject terms namely broader, narrower, and related terms as well as scope notes and Used For (UF) terms. The bottom section allows the user to view the selected query terms and the list of items retrieved by those selected terms. An evaluation of the interface carried out involving reference librarians and cataloguers has shown that the process of browsing through the thesaurus terms in Déjà vu improves users' understanding of the relationship between the archive materials and the cataloguing resources.

Sutcliffe et al. (2000b) describe a visual interface enhanced with a thesaurus called the Integrated Thesaurus-Results Browser which provides simultaneous access to query bar, thesaurus terms and structure as well as search results. The thesaurus is a major feature of the interface allowing for search term selection and query specification or modification. The thesaurus terms are shown using categorical boxes with lines connected to subcategories. Deeper levels of the thesaurus can be reached through single-clicking. Double-clicking thesaurus terms put them into the query bar which is at the top of the thesaurus area. This interface was designed based on the idea that, in the early phases of articulating needs and forming queries, users should be supported by thesauri and term suggestion facilities. A usability evaluation of the interface demonstrated that users liked the visualisations of the thesaurus and the results. The evaluation also highlighted a number of unresolved usability problems.

Blocks et al. (2002) have reported the development of a prototype search interface enhanced with the Art and Architecture Thesaurus. The interface contains a thesaurus browser for users to access the hierarchies of the thesaurus, a search facility that maps an initial search term to the vocabulary terms and a section showing the retrieved results. A formative evaluation of the interface showed that the interface was successful in allowing a person with little knowledge of the interface to make use of its functionality. However, the prototype interface did not provide non-expert searchers with sufficient guidance on query structure or when to use the thesaurus within the search process.

Commercial Web-based Interfaces

The advent of near-universal access to the Internet and the extensive use of hypertext technology gave rise to the design and development of a large number of web-based thesaurus interfaces both as standalone tools and as integral parts of database and IR systems. The focus of this section is on thesaurus-enhanced interfaces to commercial web-based bibliographic databases and their approaches to making thesauri accessible during the search process. Other types of web-based thesaurus-enhanced systems such as subject-based information gateways and multi-thesaurus browsing and searching systems are reviewed in another paper (Shiri and Revie, 2000).

There are a number of features commonly exploited in the web-based interfaces to bibliographic databases made available by commercial vendors, which characterise the thesaurus-based searching and browsing facilities that support users in query formulation and expansion. In Table 2.3 details of six interfaces are tabulated to show the extent to which these database providers have used different features for thesaurus-based search strategy support. The six bibliographic databases or providers considered are LISAnet from Bowker (now acquired by Cambridge Information Group), Cambridge Scientific Abstracts (CSA), ERIC Search Wizard, OVID technologies, PubMed from National Library of Medicine, and Silver Platter (WebSpirs). With the exception of the ERIC search wizard and PubMed, these databases are fee-based and require some kind of account for access. (The URLs of these web-based databases are listed in the references).

A number of guidelines have been suggested for comparing or evaluating interface features of thesauri (Ganzmann, 1990; Milstead, 1991). These deal primarily with thesaurus construction and editing facilities which are found in thesaurus management software. While some of these criteria

appear to be useful, there are other features that support end-user interaction which are not covered. These guidelines have therefore been augmented in the light of suggestions made in a number of the investigations reviewed earlier and an examination of the interfaces assessed here, to form the basis for devising a general framework for identification and comparison of common thesaurus interface features. Note that the features are defined based on search and browse support that they provide for end-users of the retrieval systems.

It can be seen from Table 2.3 that some of these database providers consider *thesaurus* search options as an advanced search technique, while others provide a thesaurus search bar on the main search page. The use of hypertext and hyperlink features varies from one interface to another, although almost all the interfaces use hypertext as a means for browsing and moving from a term to its thesaurus details. Some interfaces use checkboxes for selecting thesaurus terms, while others use *Add* or *Paste* buttons to place the thesaurus term in the search bar. The ERIC search wizard has applied both approaches to make the selection of search terms easier. Hierarchical and alphabetical thesaurus displays also differ in these interfaces. For instance Ovid, SilverPlatter, and LISAnet, are capable of showing both simultaneously, whereas in CSA the user must switch pages to view the alphabetical display. Most of the interfaces provide a mechanism by which the user can obtain details of narrower terms, although only PubMed and CSA use a + sign to let the user know of the existence of such terms.

Redirection from non-descriptors to descriptors is a significant facility for choosing the correct search term to query the database. Ovid performs automatic redirection, while CSA, ERIC and PubMed provide interactive redirection which means users can see both non-preferred and preferred

terms before selection. LISAnet and SilverPlatter do not support such a facility.

Receiving a prompt to inform the user that a term does not exist in the thesaurus is regarded as a useful thesaurus interface feature. Ovid, ERIC, PubMed, and CSA provide such a facility. Surprisingly, given that Pollitt et al. (1994a) demonstrated the value of providing users with information on the number of hits per descriptor, only the Ovid interface supports this feature.

Boolean operators for carrying out combined searches have been designed in different ways. PubMed, ERIC and LISA provide AND, OR, and NOT options, while Ovid and CSA allow users to select either AND or OR. SilverPlatter only supports an implied 'OR' operator.

Alternative term suggestion mechanisms based on thesaurus descriptors are also supported by ERIC, PubMed, SilverPlatter and Ovid. Based on the richness of the thesaurus, the kind of mapping or statistical techniques used, and the possible use of permuted and other types of indexes, each of these interfaces provides a number of search terms including term(s) which may not be related through the thesaurus to any of the terms entered by the user.

INTERFACE FEATURES	BOWKER LISANET	CSA	ERIC SEARCH WIZARD	OVID	PUBMED (NLM)	SILVER PLATTER WEBSPIRS
Thesaurus only in advanced search mode	Yes	No	No	Yes	Yes	No
Thesaurus button on main search page	No	Yes	No	Yes	Yes	Yes
Thesaurus search bar	Yes	Yes	Yes	No	Yes	Yes
Hypertext navigation	No	Yes	Yes	Yes	Yes	Yes
Hierarchical display	Yes	Yes	Yes	Yes	Yes	Yes
Alphabetical display	Yes	Yes	No	Yes	No	Yes
Permuted list	No	Yes	No	Yes	Yes	Yes
Scope notes	No	Yes	Yes	Yes	Yes	Yes
Explode option (including narrower terms)	No	Yes	Yes	Yes	No	Yes
Browsing previous and next thesaurus terms	Yes	No	No	Yes	No	Yes
Non-descriptor to descriptor redirection	No	Yes	Yes	Yes	Yes	No
Feedback on the term not found in thesaurus	No	Yes	Yes	Yes	Yes	No
Explicit Boolean operators	Yes	Yes	Yes	Yes	Yes	No
Display number of hits for each descriptor	No	No	No	Yes	No	No
Alternative term suggestion option	No	No	Yes	Yes	Yes	Yes

Table 2.3 Features of thesaurus-enhanced interfaces of commercial web-based databases

2.3.6 Conclusion

In this section we have discussed the rationale for incorporating thesauri into interfaces as aids to the search process and have provided an account of both research prototypes and commercial web-based interfaces enhanced through the use of thesauri. Various approaches of both a textual and graphical nature have been adopted in the design and development of prototype and research-based interfaces which incorporate thesauri. As a result the design features differ significantly from one interface to another.

Leaving to one side the most basic differences between interfaces such as the historical use of character-based command-line environments, now largely replaced by window-oriented hypertextual platforms, the key differences in the interface features present in the research-based prototypes examined earlier can be summarised as follows:

- Stage of searching at which the user is offered access to the thesaurus
- Method used to display thesaurus terms and their relationships
- Sophistication of options for browsing, selecting and deselecting thesaurus terms for querying the database
- Types of Boolean operators available during thesaurus browsing

Having observed some of the differences in approach it is also significant to note that there are a number of common features available in most of the research-based interfaces reviewed here. These include the following:

- Linking users' search terms with those in the thesauri using different mapping techniques;

- Informing the user of the number of documents indexed by a given thesaurus term;
- Look-up and browse options for finding thesaurus terms;
- A *search term pool* option for saving thesaurus terms browsed and selected by users for later use;
- Integration of thesaurus and documents displays to allow for easy query modification;
- The use of hypertext techniques for navigating and selecting thesaurus terms;
- The assumption that all types of thesaurus term relationships should be supported.

Commercial database providers tend to support a relatively uniform set of thesaurus features in their interfaces. Some of the features suggested by IR researchers have been incorporated in the design of commercial web-based interfaces while many have, to date, remained experimental techniques. From the review of the range of thesaurus-enhanced interface approaches, the facilities most likely to further improve the search process for the end user are summarised in the list below:

- An explicit thesaurus search option on the main search page in an easy-to-use way for end-users. Terms such as *suggested terms*, *thesaurus*, and *subject headings* have been used to show the availability of a thesaurus facility in the interface.
- Providing easy and understandable terminology to describe the relationships between descriptors and terms. In some interfaces term relationships are shown using such notations as NT, BT, RT, USE etc. Others have used the complete form of thesaurus relationships i.e. narrower, broader and related terms. There are also some interfaces

that have used such signs as + and - to point to the narrower or broader relationships and terms.

- Providing hierarchical, alphabetical and permuted lists to support different browsing and searching strategies.
- Flexible ways of choosing terms for posting to the search system such as drag and drop, checkboxes, hypertext features and double clicking.
- Facilitating the process and understanding of moving from a descriptor to its hierarchical structure using hypertext navigation.
- Catering for the selection of alternative Boolean operators for combining different thesaurus terms.
- Providing feedback on terms not available in the thesaurus and suggesting terms related in some way to the initial entered term.
- Providing a *term pool* option for saving the descriptors chosen by users during thesaurus browsing for later use.
- Integrating thesaurus and retrieved documents displays for more effective search and retrieval.
- Availability of thesaurus option in all stages of the search process namely query formulation, modification or expansion.

The features listed above can be implemented in a variety of ways. However, the evaluation of interfaces enhanced with thesauri and the extent to which they can aid the users' interaction behaviour must be subject to empirical research. Very few of the research-based interfaces reviewed (McMath et al., 1989; Beaulieu, 1997; Sutcliffe, 2000; Blocks et al., 2002) have been evaluated in terms of the ways in which they support query formulation and expansion, or different users' perceptions of issues such as general interaction or usability. Such evaluation

provides insights into alternative ways of designing and tailoring interfaces to meet the requirements of operational IR environments.

While the number of commercially available thesaurus-enhanced interfaces is increasing, no formal or empirical studies have been reported which evaluate the thesaurus-enhanced features of these interfaces. Research is required to evaluate the usefulness of thesaurus-enhanced interfaces in operational web-based databases with different types of end-users. The present Investigation aimed to evaluate user-thesaurus interactions in an online environment, specifically to identify interface features which are helpful to end-users and to understand more clearly the ways in which thesaurus-enhanced search interfaces affect search term selection and reformulation. Shiri and Revie report a pilot study (2001) which has proposed a methodology for the evaluation of user-thesaurus interaction in an online web-based environment. Further research provides evidence as to which interface features and facilities aid user interaction in general, and term selection, query formulation and expansion in particular.

This chapter has reviewed the literature relating to information seeking and IR interaction models, user-centred search term selection and query expansion and thesaurus-enhanced search interfaces developed in research and commercial environments. The next chapter will present the methodology and experimental environment in which the present research was undertaken.

Chapter 3: Experimental Design and Methodology

This chapter highlights the methodology and the experimental procedures developed to investigate the research questions addressed by this study. It begins by summarising general theoretical and evaluation frameworks, reviews the research questions and hypotheses, and then describes research methods, tools and techniques constituting the research environment in which the study was carried out.

3.1 Theoretical Framework

The context of the present study is the investigation of user interaction with an IR system. The particular focus is on search term selection and query expansion through the use of a thesaurus-enhanced search interface. To provide a theoretical basis for conducting this study a number of IR interaction models have been studied, the details of which appeared in the first section of the literature review. The aim of this section is to highlight those IR interaction models which have been particularly relevant in forming a theoretical basis for this study.

The three IR interaction models proposed by Belkin (1995), Ingwersen (1996) and Saracevic (1996) were taken into consideration for this study. The reason for choosing these models lies in the fact that they approach user interaction from different perspectives and each provides insights into different layers of IR interaction. Thus, for the purpose of this study elements were chosen from the three models in order to complement one another and to contribute to a better understanding of interaction.

Ingwersen's cognitive model (1996) provides a holistic view of how individual aspects of interaction can be cognitively viewed. The elements which are of particular interest within this model are: users' current cognitive state, their background knowledge, and the interface and thesauri as cognitive structures which can influence the interaction. From the polyrepresentation point of view as defined within this model, thesauri are construed as one type of representation in the cognitive space which can interact with and influence other representations such as query formulation and expansion. In addition, the notion of cognitive overlap, for instance between the cognitive structure of a thesaurus created by indexers and the users' cognitive structure, is another of the elements with relevance to the present study.

Belkin 's episode model (1995) focuses on the user's interaction with text as a central process in IR. Two elements of this model relate to information sources: information and meta-information. The model views meta-information as thesauri and classification schemes that describe the structure and contents of an information object. He identifies sixteen information seeking strategies (ISS) eight of which have to do with meta-information. Of relevance to the present study are those information seeking strategies which focus on browsing and searching using thesauri and classification schemes.

Saracevic's stratified model of interaction views the process as a dialogue between the user and the computer through an interface. The elements which are of interest in this model are those associated with the analysis and evaluation of users' interactions at the surface level: search term selection, query modifications and search strategies, as well as the user's cognitive, affective and situational levels. The relevance of this model to the present study lies in the fact that it provides an extension to search term selection which has also been applied to empirical studies (Spink and Saracevic, 1997; Spink, 2002).

As mentioned earlier each of the above models deals with interaction in a different way. However, taken together they provide an all-embracing view of the interaction with particular reference to a thesaurus-aided search term selection and query expansion environment.

3.2 Evaluation Framework

Robertson and Beaulieu (1997) have emphasised the importance of taking into account the user's interaction, perception, satisfaction and cognitive view of the search process in the course of any IR evaluation. They suggest that IR evaluation should not simply consider the issue of how good or bad a system is, rather it should contribute to an improvement in our understanding of the processes involved and their relationships in an interactive searching environment. It should also provide the basis for the design of better systems.

There have been a variety of approaches to users in IR research and evaluation. Saracevic (1995) identifies six general IR evaluation levels, each or any of which could be selected as objectives of research or evaluation.

Engineering level: hardware and software performance

Input level: input to the contents of the system

Processing level: performance of algorithms and techniques

Output level: questions about interactions with the system and the output, assessment of searching, interactions and feedback

Use and user level: questions of application to given problems and tasks

Social level: issues of impact on research, productivity and decision-making

The first three levels relate primarily to system-centred evaluation while the last three levels are particularly related to a user-centred evaluation approach and stress the importance of interaction in the IR process. Saracevic (1995) has pointed out a number of evaluation criteria used at these levels. These evaluation criteria are related to utility, success, completeness, satisfaction, worth, value, time, cost and so on. He has also commented that such criteria seem to be well suited to the evaluation of IR interactions.

Robertson and Hancock-Beaulieu (1992) refer to three major revolutions in the evaluation of IR systems: the relevance revolution, the cognitive revolution and the interactive revolution, all of which have significant implications for the involvement of users in the IR evaluation process. They note that qualitative data on users' perceptions of their interaction are important in order to evaluate IR systems. Ingwersen (2001a) views these three revolutions as the real challenge to the IR community and stresses the importance of taking users' cognitive state into account during the process of IR evaluation.

In a comprehensive review of approaches, issues and methods employed in the evaluation of IR systems Harter and Hert (1997) studied user-oriented research and identified two classes of evaluation measures: measures associated with users' perception and attitudes such as utility, usefulness, impact, satisfaction and other affective measures; and

measures whose major attempt is to capture user-system interaction. The second class of measures is derived from the HCI usability literature and includes such measures as: accuracy and completeness, error rate, process variables (e.g. number of commands, descriptors, screens accessed, search cycles), training time, retrieval (e.g. number of entries retrieved, unit cost), perceptions of ease of use, satisfaction, and ability to articulate system models.

Sugar (1995) has reviewed the user-centred perspective of IR research methods and has divided user-centred IR research into two main categories: cognitive and holistic. The cognitive approach focuses on discovering cognitive characteristics of the user's information seeking behaviour. The holistic approach concentrates not only on the cognitive aspects of an information search but also on the affective and physical aspects such as user satisfaction and perception and the use of system features.

In the present study a user-centred approach to the evaluation of users' interaction with an IR system has been adopted. It involves elements from both approaches identified by Sugar (1995), namely cognitive and holistic. Firstly, cognitive behaviours and human information processing tasks contributing to the search process have been studied. Secondly, affective aspects of searches have also been taken into consideration, in particular users' attitudes and perceptions toward the system.

The reason for choosing the holistic approach is due to the fact that the present study investigates cognitive, physical, and affective aspects of users' search behaviour. In particular, it takes account of elements such as users' perceptions, attitudes, satisfaction and impressions when evaluating users' behaviour and interaction with the system. It is holistic in the sense that involves users' cognitive and physical behaviour throughout the search process.

The general evaluation framework adopted in this study incorporates a range of user-oriented measures identified by Harter and Hert (1997):

- Measures related to perception and attitude: user satisfaction, impact and users' perception of and attitude toward the system;
- Measures mainly associated with usability: search process measures, retrieval, search time, ease of use, and ease of learning.

The first category of measures deals with affective behaviour, for instance whether or not users are satisfied with their search terms or search results, or what is their impression and perception of the search process. The second category of measures provides evaluation criteria for users' interaction with the system. For instance, how usable and learnable a system is with respect to other variables such as users' search tasks, time, screen accessed and the number of commands used.

The rationale for adopting the above evaluation framework lies in the fact that the present study intended to investigate the search process and the associated variables as well as users' interaction with the system. Therefore, evaluation measures that are not relevance oriented but place more emphasis on the user and the search process are of particular importance in this context. These measures include usability, usefulness, user satisfaction and perception.

3.3 Research Questions, Variables and Hypotheses

The primary goal of the present investigation is to evaluate how and to what extent a thesaurus-enhanced search interface assists end-users in selecting search terms for query expansion. Specifically, it attempts to ascertain users' attitudes toward both the thesaurus and interface as tools to facilitate search term selection for query expansion. It also identifies searching and browsing behaviours of users interacting with a thesaurus-

enhanced interface attached to a large bibliographic database available on the Internet.

3.3.1 Research Questions

The following research questions were investigated:

- Are there common patterns of user behaviour in thesaurus-based browsing and searching?
- What relationships are there between users' initial query terms and the terms they select from the thesaurus for query expansion?
- Does topic complexity affect user-thesaurus interaction in general and search term selection in particular?
- Does topic familiarity affect user-thesaurus interaction?
- Does interface usability affect thesaurus browsing or other search behaviours?

3.3.2 Variables

The choice of variables and level of analysis in studies on user-system interaction and searching behaviour depends on the purpose of evaluation and the research question (Borgman et al., 1996). Tague-Sutcliffe (1992) in an overview of pragmatic issues in IR experimentation describes a number of variables in IR. These relate to the following: the database, information representation, users, queries, search intermediaries, retrieval process, and retrieval evaluation. A number of information searching and retrieving studies have chosen variables such as search process, search outcome, search tasks, or searcher and user characteristics (Fenichel, 1981; Howard, 1982; Saracevic and Kantor, 1988; Marchionini, 1989; Spink and Saracevic, 1997; Efthimiadis, 2000). Other investigations have considered search tactics and strategies, search experience and subject knowledge as major variables when studying the search behaviour of different user groups (Hsieh-Yee, 1993; Solomon, 1993). Other types of variables relating to interface features

have also been reported in the literature (Meadow et al., 1995; Beaulieu, 1997; Park, 2000).

The variables defined and studied in the present investigation are based on the key research questions and are categorised as follows:

- User characteristics (academic status, subject knowledge, search and computer experience)
- Search topic characteristics (topic complexity, topic familiarity, search type)
- Search process characteristics (cognitive and physical moves)
- Search term characteristics (number and type of search terms browsed, selected, and expanded, usefulness of thesaurus terms)
- Thesaurus interface characteristics (thesaurus browsing/navigation, ease of use and ease of learning)
- User satisfaction (search terms, search results, result-topic match)

A more detailed description of the variables associated with the search process and topic complexity and familiarity are discussed in Section 3.7.

3.3.3 Research Hypotheses

A set of main and sub-hypotheses was developed in order to test relations between the defined variables and so increase knowledge relating to the general research questions. The study investigates five general hypotheses and a number of sub-hypotheses associated with each main hypothesis. The main and sub-hypotheses defined are as follows:

Main hypothesis 1: User characteristics have an effect on searching and browsing behaviour.

$H_{1\text{-}1}$: Users with more subject knowledge (i.e. academic staff) make more cognitive and physical moves.

$H_{1\text{-}2}$: Academic staff and postgraduate researchers differ in the number of terms they browse and select.

$H_{1\text{-}3}$: There is no significant difference in cognitive and physical moves made by male and female users.

$H_{1\text{-}4}$: Level of computer experience has an effect on users' cognitive and physical moves.

Main hypothesis 2: Topic complexity affects user-thesaurus interaction.

$H_{2\text{-}1}$: Increased topic complexity leads to an increase in the number of cognitive and physical moves.

$H_{2\text{-}2}$: Searches with complex topics require more query reformulation than those with simple topics.

$H_{2\text{-}3}$: Searches with complex topics are associated with the selection of more search terms.

Main hypothesis 3: Topic familiarity and search type have an effect on searching and browsing behaviour.

$H_{3\text{-}1}$: Topic familiarity, topic search experience and search type affect cognitive and physical moves.

$H_{3\text{-}2}$: Topic familiarity affects the number of search terms browsed and selected.

$H_{3\text{-}3}$: Topic search experience has an effect on the users' judgment of additional terms provided by the thesaurus.

Main hypothesis 4: Search process characteristics affect search term and search result characteristics.

$H_{4\text{-}1}$: There is a relationship between result relevance and cognitive and physical moves.

$H_{4\text{-}2}$: Cognitive and physical moves affect users' satisfaction with the number of records retrieved.

$H_{4\text{-}3}$: There is a relationship between cognitive and physical moves and users' evaluation of term usefulness.

$H_{4\text{-}4}$: There is a relationship between the number of selected terms and users' evaluation of term usefulness.

$H_{4\text{-}5}$: There is a relationship between the number of selected terms and users' judgement of the relevance of result sets.

$H_{4\text{-}6}$: Users who spend more time looking at search terms make more cognitive moves.

Main hypothesis 5: Perceptions of interface usability have an effect on searching and browsing behaviour.

$H_{5\text{-}1}$: There is a significant relationship between the ease of learning and using the interface and ease of thesaurus browsing.

$H_{5\text{-}2}$: There is a significant relationship between ease of thesaurus browsing/navigation and cognitive and physical moves.

$H_{5\text{-}3}$: There is a significant relationship between ease of thesaurus browsing/navigation and the number of terms browsed and selected.

$H_{5\text{-}4}$: There is a significant relationship between ease of thesaurus browsing/navigation and the number of query reformulation instances.

$H_{5\text{-}5}$: There is a significant relationship between the ease of thesaurus browsing/navigation and the time spent per search.

3.4 Experimental Design

The following sections outline the components of the experimental design developed within this study:

- The system
- The participants
- Search environment
- Search requests

3.4.1 System

The choice of laboratory or operational IR system depends on the objectives of the research. Robertson and Hancock-Beaulieu (1992) have stressed the importance of research within operational settings in order to encompass the human element of the IR process. For this study a large operational IR system was employed. The CAB Abstracts database, one of the largest bibliographic databases in the agricultural sciences, was utilised as the experimental platform. It covers the research and development literature in the fields of agriculture, forestry, aspects of human health, human nutrition, animal health and the management and conservation of natural resources. In 2002 it contained more than four million records. The CAB Abstracts database was chosen due to the fact that the interests of the research population were well covered by this database. The assumption was that the CAB thesaurus represents a well-developed manually constructed thesaurus which has been used by indexers and professional searchers within the context of an information retrieval system. The database is made available through various vendors including Silver Platter, Dialog and Ovid, as well as the CABI organisation itself.

The Ovid version of CAB Abstracts on the Web was selected for this study as it served the purpose and requirements of the present investigation. The Ovid search interface accommodates a range of thesaurus-enhanced facilities in its advanced search mode which allow users to map, browse and search their terms. The terms entered by the user are mapped to the CAB thesaurus terms through an interface facility labelled "Map Term to Subject Heading" (Figure 3.1). The CAB Thesaurus is the largest agricultural thesaurus and contains about 59,000 terms, of which 48,500 are *preferred terms* (descriptors) and 10,500 are *non-preferred terms*. The thesaurus provides all the standard thesaurus relationships and features. The reason for choosing the CAB thesaurus is

because the researcher is well familiar with the thesaurus and is also aware that it is a long-standing agricultural vocabulary which has been developed and used for indexing and searching over the last two decades. The interface makes use of the thesaurus to provide alternative mapping types. Depending on the users' initial query,

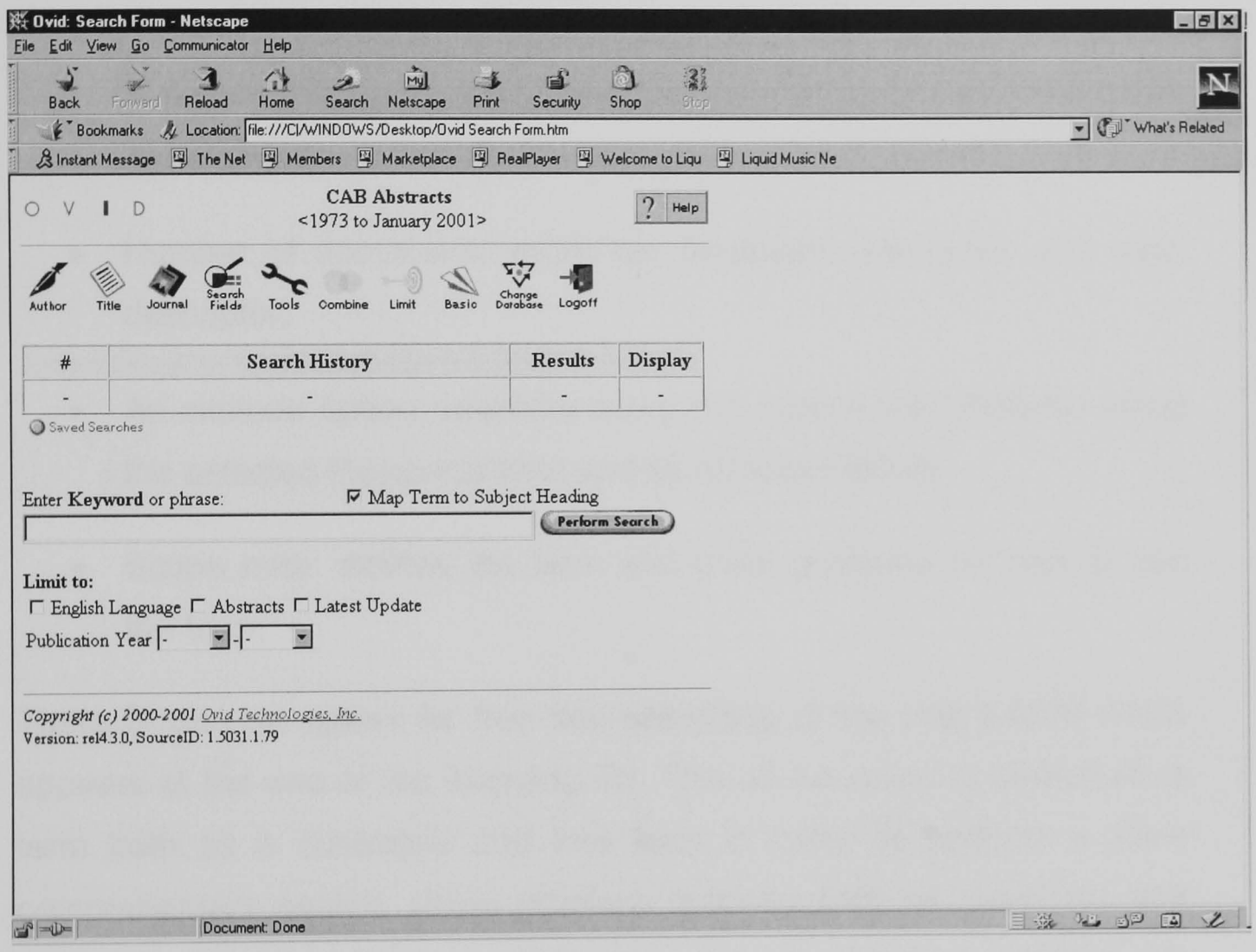


Figure 3.1 OVID main search page

the interface shows one of the following three match types to users who may then browse and select terms for querying the CAB database:

- Exact match: the user's term matches exactly to one of the thesaurus terms;
- Partial match: the user's term matched partially to one of the thesaurus descriptors which are arranged in a permuted index;

- Statistical match: through a statistical analysis the descriptors frequently co-occurred with the terms supplied by the user are shown. This technique determines descriptors that occur in documents containing the free text query.

When a term is mapped to the thesaurus the following options will be displayed to the user:

- A hierarchical list of thesaurus terms with hyperlinks to the details of a term together with a check box next to each term to facilitate its selection;
- Number of documents within the database associated with each descriptor;
- An explode option: whereby users can search the database using the selected thesaurus term and its narrower terms;
- Scope note: defines the term and gives guidance on how to use the term.

There is also an option for free text searching of the user's term which appears at the end of the mapping list. This allows users to search for a term both as a controlled and free term in order to perform a more comprehensive search. Using interface features such as hyperlinks and different graphical displays users are able to select descriptors for searching the database or click on hypertextual descriptors to view term relationships in the thesaurus and to eventually submit the query to the database (Figure 3.2 shows a typical OVID thesaurus screenshot).

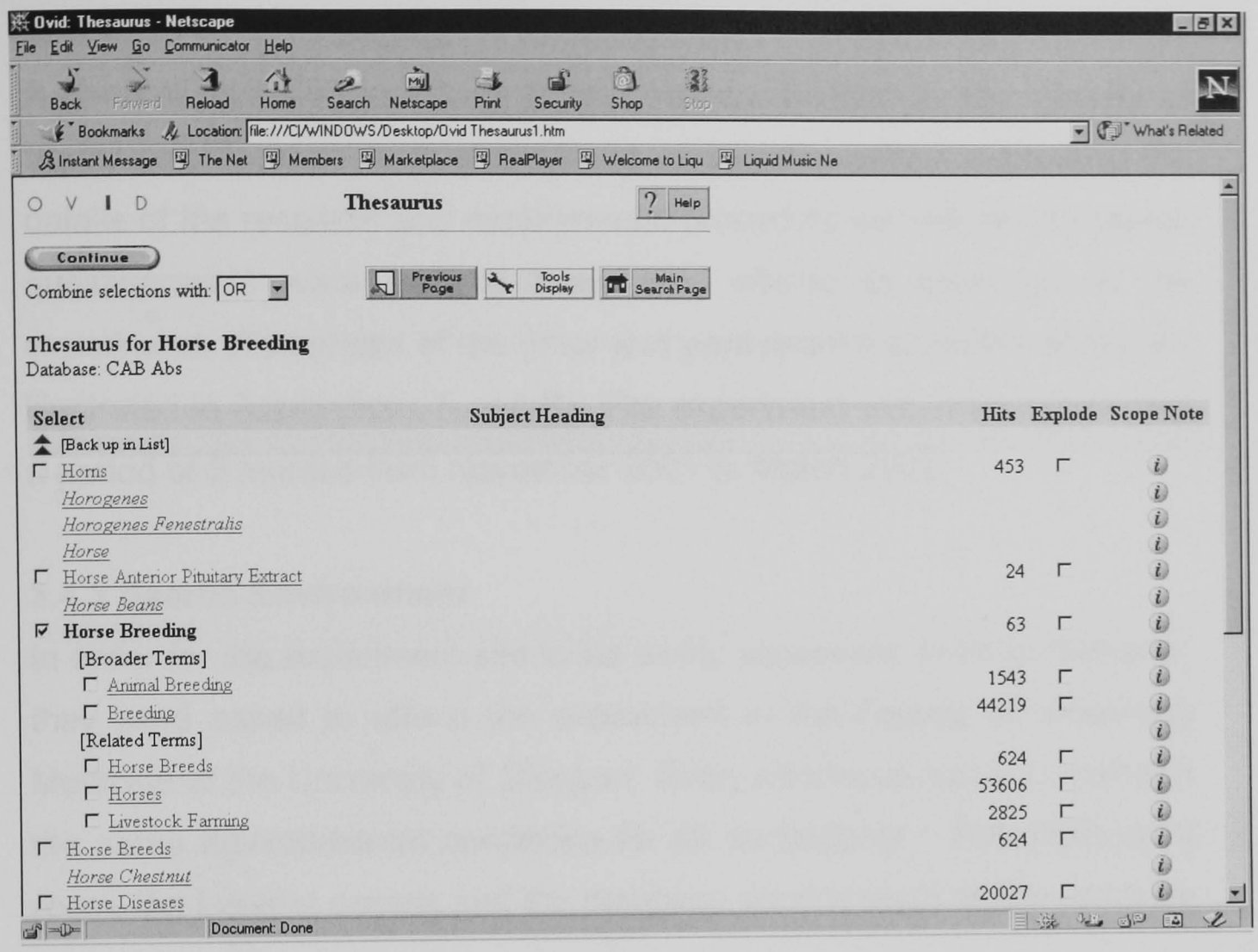


Figure 3.2. OVID thesaurus interface

3.4.2 Participants

The purpose of this study was to explore the term selection and thesaurus interaction behaviour of real users with genuine information needs. To this end, it was decided to involve researchers from a particular knowledge domain. The assumption was that academic staff and postgraduates represent two different levels of subject knowledge. The participants in the experiment were 30 academic staff and postgraduate researchers, who were selected from the Faculty of Veterinary Medicine at the University of Glasgow. The reason for selecting this subject area lies in the fact that it was well-covered by the CAB Abstracts database. Fifteen users were academic staff and the rest were postgraduates and doctoral students. Most students were *clinical scholars* which means they were involved in veterinary practice as well as research.

A general invitation was posted to a listserv hosted by the Faculty of Veterinary Medicine inviting people to participate. Letters explaining the details of the research and experimental procedure as well as pre-search questionnaires were sent to those who offered to take part in the experiment. (Templates of the letter and participant's summary sheet are illustrated in Appendices A and B). The experiment was undertaken over a period of 5 months from November 2001 to March 2002.

3.4.3 Search Environment

In order for the experiment site to be easily accessible to all participants, they were asked to attend the experiment in the Faculty of Veterinary Medicine at the University of Glasgow. Every effort was made to maintain the same environmental conditions for all participants. The experiment room, the Internet access and the database version were all the same for all participants.

3.4.4 Search Requests

There are two main types of search task reported in the literature which can be used depending on the aims and objectives of the study: assigned search tasks and topics which are usually defined by the researcher (Marchionini et al., 1991; Hsieh-Yee, 1993; Brajnik et al., 1996, Sutcliffe, 2000a) and real search requests and topics which are those elicited from users in the light of their actual information needs (Jones et al. 1995; Efthimiadis, 2000; Greenberg, 2001). Borland and Ingwersen (1997) introduced the concept of the simulated information need situation. It is characterised by a request, a definition of the request and a simulated work task situation.

In the present study, search topics provided by the participants were used to create a situation in which participants could have natural interaction with the thesaurus and the interface. This decision was made on the

assumption that the evaluation of search term selection, query expansion and users' interaction with the thesaurus can be more effectively carried out if users with genuine queries are involved in the study. The subjects were asked to provide, in the pre-search questionnaire, three search topics on which they would like to conduct searches. They were also asked whether or not they have previously carried out searches on the topic. (A discussion of the effect of prior topic search experience on users' search behaviour is provided in section 6.4 of the discussion Chapter).

3.5 Data Gathering Techniques

In a discussion of methodological issues concerning the evaluation of interactive IR systems, Robertson and Hancock-Beaulieu (1992) noted that evaluation of the search process calls for both quantitative data on user interaction with the system and qualitative interpretations of user behaviour to characterise complex behavioural processes in online searching.

The complex nature of evaluating users' behaviour and their interaction with an IR system calls for a combination of data gathering techniques to ensure an understanding of the whole process. To collect data on a wide range of variables this study employed both quantitative and qualitative techniques. The data gathering tools utilised were pre-search questionnaires, screen-capturing software, post-search questionnaire and post-session interviews.

3.5.1 Pre-search Questionnaires

The pre-search questionnaire was designed to provide data on two types of variables, user characteristics and search request characteristics (A template of the questionnaire is illustrated in Appendix D). The data that were sought through the questionnaire included: personal information,

computer and search experience, search intention, search topics and search types. The questionnaire asked for the user's age, gender, academic status and educational background, their experience in searching CD-ROM, online, and Internet databases as well as their experience in using specific computer applications. The participants were also asked to provide as part of the pre-search questionnaire three search topics based on real information needs for which they intended to perform searches. There were questions concerning the type of use they would make of the information they were searching for. For instance for MSc or PhD thesis, research project, academic publication or teaching purposes.

The users' familiarity with their search topics was recorded based on a three-point scale: unfamiliar, moderately familiar, and very familiar. They were asked whether they had any previous search experience on those topics and if they did, how long ago the search was conducted. To identify the type of search users intended to perform, they also answered a question on whether they wanted to carry out a *broad* or *specific* search.

3.5.2 Transaction Logs

To collect data on all aspects of user-system interaction, this study adopted an online monitoring technique. Online monitoring, also known as transaction log analysis, is a popular and proven technique for evaluating searching behaviour and complements other methods such as interviews and visual observations (Borgman et al., 1996). It is also considered to be a powerful and flexible technique for data collection in order to analyse user interactions with an IR system and to evaluate IR interfaces (Penniman and Dominick, 1980). Jones et al. (1997) state that perhaps the most important purpose of logging is to provide material for analysis. It helps experimenters to form a complete picture of an interactive search session.

Within this study Lotus ScreenCam, a commercial screen-capturing software package produced by IBM Software Group, was utilised. It records all actions taking place on the screen and creates a movie file for each search session for later replay. It also provides such features as length, contents and revision history of the movie. This online monitoring technique collected data on variables associated with search process characteristics, search term selection characteristics, and thesaurus interface characteristics. The screen-capturing software gathered data on the entire session of each search performed within the Ovid CAB Abstracts database. Data collected through this technique comprised all moves and actions users undertook during the search process including the selection of search terms derived from search topics, thesaurus mapped terms, different thesaurus displays, browsing and navigation around terms, use of individual system features, number and types of terms browsed and selected, search reformulation, and search results. The screenshots that were used for analysis are provided in Appendix J. It should be noted that the transaction logs were stored anonymously with no reference to the individual who created the log, and that consideration has been given to how the video-captured data will be disposed.

3.5.3 Post-search Questionnaires

A post–search questionnaire was designed to collect data on variables connected with search term and search result satisfaction (A template of the questionnaire is illustrated in Appendix E). The participants were asked whether the thesaurus provided them with additional search terms and whether or not they were useful. A judgement on the closeness of the terms provided by the thesaurus to the original query and an indication as to whether the user was aware of the additional terms at the beginning of the search, were also elicited. Reasons for the selection of additional terms and the participants' satisfaction with the number of search terms were also sought. The second part of the post-search questionnaire

aimed to gather data on users' satisfaction with search results and the extent to which the results matched their original query. The participants were asked to rate the relevancy of the result set for each search based on a three-point scale: relevant, partially relevant, not relevant. This decision was made based on the assumption that users' assessment of result sets as opposed to individual documents provides an easier way of evaluating users' impressions and satisfaction. The Text REtrieval Conference (TREC) Interactive Track has also recently carried out relevance assessments based on document collections rather than individual documents (Guidelines for TREC –2003 Interactive Track).

3.5.4 Post-session Interviews

Post-session semi-structured interviews were carried out to collect data on the usability of the interface and the thesaurus (A template of the interview script is illustrated in Appendix F). Participants were asked to rate the ease of use, ease of leaning and ease of browsing and navigation based on a five-point scale. They were also asked about any difficulties they encountered while interacting with the interface or the thesaurus and the type of help they required throughout the search process. In the interview users were asked to comment on the effect of browsing on their search term selection and the features they liked or disliked. Users' general perceptions on their interaction with a thesaurus-enhanced search interface were also sought during the interview.

3.6 Experimental Procedure

The following provides the sequence of stages users passed through during the experiment:

1. Agree to take part by e-mail
2. Fill in the pre-search questionnaire
3. Receive a short introduction to the experiment and the system
4. Conduct a practice search

5. For each of three pre-specified topics carry out a search as follows:
   - 5.1 Enter candidate search terms
   - 5.2 Work through different thesaurus-based search stages including browsing and selecting terms from mapped thesaurus descriptors and terms
   - 5.3 Possible search reformulation
   - 5.4 View search results
   - 5.5 Fill in the post-search questionnaire
6. Participate in a post-session interview

Pre-search questionnaires were completed and returned by participants approximately one week before the experiment was carried out. Participants came one at a time to attend the experiment. Each participant was initially given a brief account of the purpose of the study and the procedure of the experiment. To ensure the consistency of users' introduction to the system, they were asked to read a short written tutorial on performing searches using the Ovid CAB Abstracts database. (The written tutorial is illustrated in Appendix C). The tutorial provided users with a short description of the database, the search facilities and in particular the thesaurus mapping facility which was available in the advanced search mode. Having read this each participant was asked to conduct an assigned practice search to familiarise themselves with thesaurus mapping and other system features. The practice search was the same for all participants. The practice search was designed in such a way as to provide the users with an opportunity to gain experience of a typical thesaurus-based search session and to demonstrate the manner in which thesaurus terms and structure were presented. Since the OVID thesaurus mapping facility can only effectively operate using individual concepts, users were instructed to adopt a *building blocks search strategy*. This strategy requires that a search request is split into single concepts and that one concept is searched at a time along with its all narrower, broader, related or synonymous terms before all concepts are

finally combined and used to conduct the entire search. Users were asked to decompose their search topic based on the individual concepts that they felt constituted each topic. The rationale of this procedure lies in the results of the pilot experiment carried out by the researcher (Shiri and Revie, 2001).

Users were then asked to perform their three searches one after the other using the search strategy mentioned earlier. They were also told that they would be interrupted after each search to fill in the post-search questionnaire. In total 90 searches were conducted by 30 users.

The participants conducted their searches using the thesaurus mapping facility in the Ovid advanced search mode. Terms entered by users were automatically mapped by the Ovid search interface to thesaurus descriptors and the users were provided with a list of suggested terms. They could choose various types of thesaurus terms using hyperlinks and check box options. The *explode* option and free text searching of thesaurus descriptors were also available to users. After selecting the terms, users would post all selected search terms to retrieve the records. After viewing the results, the user could send the whole or part of the retrieved set to their own e-mail address using an option available in the Ovid interface. The entire search session of each user was captured using the Lotus ScreenCam software. The user was then asked to fill in the post-search questionnaire. This search procedure was followed for each of the user's three search topics.

After completing the whole search process, a semi-structured interview, which was audio-taped, was carried out with each user. Users were asked to comment on issues related to general usability, the usefulness of the thesaurus and their general impression on the whole search process. On average it took each participant one and a half hours to complete the experiment including searching for the three topics, filling in

the post-search questionnaires, and the post-session interview. Table 3.1 summarises the quantitative and qualitative data gathering tools utilised in this study.

Quantitative data	• Pre-search questionnaire • Transaction log • Post-search questionnaire
Qualitative data	• Post-session interview

Table 3.1 Data gathering tools

In this study quantitative data was defined as those data elements which were analysed using basic or inferential statistics such as the number of cognitive and physical moves, complex and simple topics and broad or specific searches. Qualitative data, in contrast, was defined as those data elements which have been subject to content analysis or have been used to extract concepts, categories of thoughts or ideas or descriptive information provided by users, for instance the problems users experienced when interacting the system.

The experimental design and data gathering techniques employed in the study have been expanded from a pilot study (Shiri and Revie, 2001; Shiri et al., 2002c) with four faculty researchers undertaken in November 2000 to ensure the feasibility of the experiment on a small scale. The initial data gathering tools included a pre-search questionnaire, screen capturing software and post-session interview. The post-search questionnaire was then developed for the full-scale experiment in order to gather more detailed information on each particular search session. This caused some changes in the experimental procedure as users had to be interrupted after each search session to fill in the post-search questionnaire.

3.7 Data Analysis Methods

Quantitative and qualitative data gathered through pre-search questionnaires, screen capturing software, post-search questionnaires and post search interviews were analysed at two levels namely the user level and the search level. The rationale for defining these two levels of analysis was due to the fact that the thirty users in the study conducted ninety searches, three searches each. Therefore, an analysis of individual users and searches was essential. Variables such as status, gender and computer experience were analysed at the user level. For other variables such as topic complexity, topic familiarity and type of search the analysis was conducted at the search level.

In order to organise all the quantitative and qualitative data for proper analysis, a Microsoft Access database was developed and all the gathered data were input. This practice was particularly helpful since the data could easily be exported to other applications such as Microsoft Excel and Minitab for basic calculations and statistical analysis respectively. The methods used to analyse data on search behaviour, topic complexity and topic familiarity are described in the following.

3.7.1 Search Process Measures

To identify thesaurus-based search behaviours and patterns a number of search process measures were defined to analyse all individual search process characteristics. These measures were search state, search term, and search move.

Search state: major stages or conditions of a system that a user goes through while conducting a search. Five main states were defined as characterising a typical end-user search interaction with a thesaurus-enabled search interface. Figure 3.3 shows the definition and sequence of the search states.

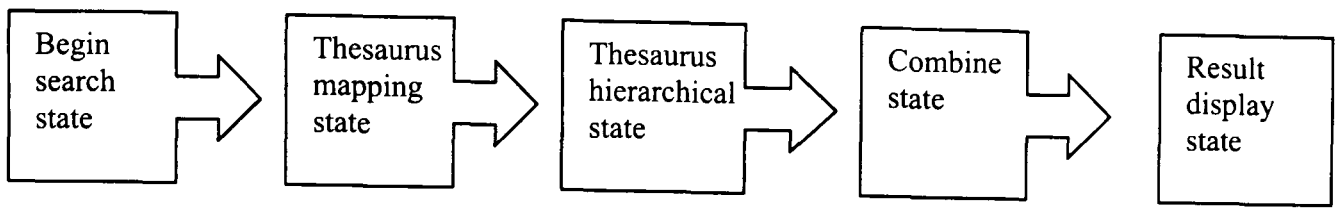


Figure 3.3 Thesaurus-based search states

Search term: a general characterisation of all types of terms during the entire search, provided either by the user or the system. The types of search terms viewed or selected were analysed and categorised as follows:

- initial search terms provided by users
- terms mapped to the thesaurus descriptors (exact, partial or statistical match)
- terms from the hierarchical structure of the thesaurus namely narrower, broader and related terms
- query expansion terms

The frequency of the number of all terms browsed and/or selected by users at various interaction levels were counted and tabulated to serve the purpose of statistical analysis.

Search move: characterises any action a user takes while interacting with the system. Marchionini et al. (1991) in studying user interaction with IR interfaces have defined a move as a conceptual act manifested as one or more keystrokes for instance entering an entire query or pressing the page down key. This definition has been adopted in this study. Various approaches have been taken in defining search moves in studying online search behaviour. For instance Bates (1979) has introduced a number of search moves or tactics some of which specifically relate to search formulation and search terms. Fidel (1985) has identified two types of online search moves namely operational moves and conceptual moves.

These moves are more specific than Bates' search tactics in the sense that they are mainly concerned with the use of free-text and controlled vocabulary terms for broadening or narrowing down the search. However these two sets of moves were primarily suggested for professional online searchers rather than end-users. Solomon (1993) has studied children's information retrieval behaviour and has proposed a number of moves. These are: opening moves, simple (single concept) and complex moves (multiple concepts), exploratory moves to investigate the system's features, and confused moves where the user was not able to perform basic tasks using the system. In a study of end-user searching behaviour Manglano et al. (1998) classified search moves based on the following stages: choose search system, formulate query, query reformulation, examine results, extract information, and stop.

Since the theoretical framework underlying the present study has focused on cognitive aspects of user interaction with IR systems two general categories of moves were defined for the purpose of analysis namely *cognitive moves* and *physical moves*.

- Cognitive moves are those in which users perform some kind of conceptual analysis of terms and/or documents;
- Physical moves are those associated with the use of system features

These move types were considered to be appropriate for the present study because:

- they represent actions and steps users might take within a thesaurus-based searching and browsing environment
- they provide various levels of granularity for the analysis
- they reflect some move types which have been adopted in previous studies (Marchionini, 1989; Marchionini et al., 1991; Mangelano, 1998)

To create a framework for capturing all actions and moves made by users, a number of ScreenCam files were randomly selected to analyse and identify the types of moves throughout the whole search process. Each video was replayed first and all actions, steps, moves and mouse movements were recorded. This was an evolving process because not all users made the same use of system features. Based on the recorded actions and search states two categories of moves were identified. These move types were associated with the variables *thesaurus interface* and *search interface*.

Table 3.2 and 3.3 show the two categories of moves identified and a description of each move type.

Cognitive moves	Description
Term input	input initial concept
Browsing terms in mapping state	browsing system suggested terms in the mapping state
Browsing terms in thesaurus hierarchy state	browsing system suggested terms in the hierarchical thesaurus state
Selection of terms	selection of terms in either mapping or hierarchical state
Combine search terms (Boolean)	choosing to combine search terms
Browsing retrieved titles	browsing the titles of retrieval records
Query reformulation	any move to modify, enhance or expand the query

Table 3.2 Description of cognitive moves

Users can browse thesaurus terms in two ways. After a user inputs a term a list of exact, partial, or statistical match is displayed. This allows users either to browse the list or click on a particular thesaurus term to see the details of the term in hierarchical mode. Therefore, two types of browsing have been identified in the categorisation of cognitive moves namely *browsing in mapping state* and *browsing in thesaurus hierarchy state*. *Combine search terms* indicates the number of times users made a decision as to which search terms would be combined using Boolean

operators. As a general cognitive move *query reformulation* involves any effort to modify, enhance or expand the query.

Physical moves	Description
Perform search	to send the query
Scroll up & down	actual number of mouse movement of the scrollbar
Back & forward	the browser option for backtracking or moving forward
Continue	to move forward in every stage of the search
Combine	clicking combine button
Citation display and e-mail	option for choosing to view and e-mail the retrieved set
Main search page	returning to the main search page
Previous & next page	buttons for flicking through result pages
Search history	choosing to view the search history
Expand & contract	buttons for collapsing or contracting search history

Table 3.3 Description of physical moves

The *Back and forward* option in the list of physical moves refer to the browser's facility for backtracking or going forward, while *Previous and next page* are facilities for flipping through result pages. The *Continue* feature allows users to proceed with the search process on the mapping page, the thesaurus page and the combine search page. The *expand and contract* option cater for collapsing or contracting the list of previous search statements in the main search page and the combine search page. The *Combine* category denotes the number of times the user clicked the icon which activated one of the Boolean operators.

The process of analysing and counting the number of all cognitive and physical moves made by users over ninety searches was time consuming as the researcher had to replay each ScreenCam file a number of times to ensure that all identified moves had been analysed and recorded at both user and search levels.

To record time spent on searches and the number of screenshots users went through, three levels were defined: (i) time and number of

screenshots per concept; (ii) time and number of screenshots for combining search terms and viewing results; and (iii) time and the number screenshots per search.

It should be noted that in this study all cognitive and physical moves were treated as being equal in terms of their importance since all moves were counted and analysed without any reference to their weight or role in the search process.

3.7.2 Topic Complexity

One of the search topic characteristics examined in this study was topic complexity. Borgman et al. (1996) have pointed out that search topic remains useful as an independent variable but it is more useful when used to investigate the effect on search process variables rather than on product variables.

Various approaches have been taken to study and identify the search task and its characteristics. These can be generally categorized as search task characteristics, query structure characteristics, and search topic characteristics. Hansen (1999) has identified a list of search task characteristics and their importance in designing an interface for IR interaction. Some of the characteristics include: simple versus complex, active versus passive, structured versus unstructured, and so forth. Task complexity and its effect on information seeking has also been investigated and it has been shown that task complexity has systematic and logical relationships with types of information, information channels and sources (Bystrom and Jarvelin, 1995). Research has been carried out on query structure and query complexity and their effects on information retrieval. Kelalainen and Jarvelin (1998) have defined query structure as using operators to express the relations between search keys and have classified queries into weak and strong structures. They have found that strong query structures lead to better retrieval

performance. In a study of the effect of query complexity on web searching results Jansen (2000) has used the simple versus complex dichotomy based on the number of search terms per query and the Boolean operators used. He concluded that increasing the complexity of a query had little effect on search results. Koenemann et al. (1994) have characterized queries by three factors: the use of terms in the query, the use of operators, and the use of relevance feedback to expand the users' formulated query and to reweight the terms.

Saracevic et al. (1988) describe five categories of questions based on domain, clarity, specificity, complexity and presupposition. They define complexity based on the number of search concepts. Their study concluded that questions with higher complexity had an increased probability of retrieving more relevant and precise results. Efthimiadis (2000) has investigated request characteristics within the context of an interactive query expansion environment. The criteria he used included the subject area of the request, nature of enquiry (accurate or vague), and type of search required (broad or narrow). However, these variables have not been investigated with respect to users' search behaviour. Fowkes and Beaulieu (2000) in an investigation of searching behaviour in a highly interactive environment classified simple and complex topics based on the level of engagement of users with the content and the effort required to find the answer to the topic. They found that effectiveness of relevance feedback was dependent on topic characteristics. They concluded that automatic query expansion is more useful for simple topics while interactive query expansion with contributions from both the searcher and the system appeared to be more effective for complex topics.

In this study the analysis of topic complexity has been based on two criteria: the number of search terms per each topic and the number of Boolean operators used for changing a topic into a query. The number of

search terms was counted for each topic to provide a quantitative perspective on all search topics. The number of AND and OR operators used in formulating queries based on the topics was also counted. These two criteria were used to categorise topics as either simple or complex. Topics having more than three search terms or more than two operators were classified as complex while those with three search terms or less and two operators or less were categorised as simple topics. The motivation behind this choice of complexity measure lies in the fact that a) both initial search statements and Boolean operators are important elements in the process of transforming a topic into a query; and b) previous research (Saracevic et al., 1988; Koenemann et al., 1994; Kelalainen and Jarvelin, 1998; Jansen, 2000) has indicated that the number of search terms and operators play a significant role in defining the complexity level of a topic. This classification led to around 60% of the topics being defined as simple.

3.7.3 Topic Familiarity

Several studies have investigated users' familiarity with search topics and its effect on search behaviour. Hsieh-Yee (1993) studied the effect of subject knowledge and search experience on users' search tactics. She found that subject knowledge affected the tactics experienced searchers adopted. Experienced searchers who were not familiar with the topic relied more on the thesaurus for term suggestion, made more effort in preparing for the search, included more synonyms and tried more combinations of terms whereas experienced searchers familiar with the topic used more of their own terms.

Vakkari (2002) studied subject knowledge, source of terms, and term selection in query expansion and suggested that users with more subject knowledge have a larger vocabulary and tend to identify more variant expressions of a concept. Kelly and Cool (2002) investigated the effects of topic familiarity on information search behaviour. They concluded that

some information search behaviours like reading time and efficacy, which they measured as the ratio of the number of saved documents to the total number of documents viewed, vary with respect to topic familiarity. They indicated that the general trend suggests that the more familiar a user is with a search topic, the less time they spend reading.

The present study evaluated the users' topic familiarity based on a three-point scale: very familiar, moderately familiar and unfamiliar. The users were also asked whether or not they had previously searched for each topic and, if so, the length of time since their last search.

3.7.4 Verbal Data

The verbal data gathered within this study includes opinions, impressions, comments and critiques made by users about their interaction, search experience, the problems and difficulties they faced, and usability issues. Five questions in the post-session interview and one question in the post-search questionnaire sought these data. The content analysis approach was adopted in order to explore patterns, themes and categories in the data. As stated by Patton (2002) content analysis is used to refer to any qualitative data reduction and sense-making effort that takes a volume of qualitative material and attempts to identify core competencies and meanings. In this study the responses were analysed at question level and the major categories, themes and concepts of each response and their distribution were recorded. The concepts were then categorised as appropriate to each question. For instance a question was asked about the effect of thesaurus browsing on users' search term selection. Users employed different words to describe the effect of thesaurus browsing using such terms as *narrow down*, *broaden* or *get more specific* and *widen*. These terms were categorized and their distribution was counted to provide a basis for general impressions that the users had of the thesaurus, the search interface and their general search experience.

The result of verbal data analysis was also compared with other variables to find relationships between them.

3.7.5 Statistical Data Analysis

Descriptive as well as inferential statistical methods were used to analyse the data. The initial step was to provide such descriptive statistics as frequency counts and cross-tabulations. To analyse the relationship between the variables in the hypotheses listed above, a number of statistical techniques were used. Since the variables were both continuous and ordinal, different statistical tests were run. Mean, standard deviation and correlation coefficients were initially calculated for all cognitive and physical moves as well as all other variables.

Ordinal variables such as search term satisfaction, search result satisfaction and relevance of retrieved records were all numerically coded to allow for the statistical testing. A *t* test was used to compare variables with two groups. Variables such as search type (broad and specific) and the type of users (academic staff and postgraduates) were therefore *t*-tested. For variables with more than two groups, analysis of variance (ANOVA) was used. Within this category fall such variables as topic familiarity and its effect on cognitive and physical moves. For tests of significance on relationships between categorical variables, *Chi-square* tests were conducted.

One of the main issues in statistical data analysis is how to treat 'outliers'. Outliers are defined as extreme cases on one variable, or combination of variables, which have a strong influence on the calculation of statistics. In the process of statistical data analysis for this study 4 and 2 outliers were identified for cognitive and physical moves respectively and these have been removed where appropriate to ensure the normalisation of the data.

3.8 Summary

This chapter has described the methodology and the experimental environment within which the present study was conducted. Research design including research questions, variables and hypotheses has been presented. Data collection tools, techniques and procedures were discussed. The statistical and verbal data analysis methods used were also introduced. The next chapter presents the descriptive (qualitative) findings of the experiment reported in this chapter.

Chapter 4: Descriptive Data Analysis

This chapter presents descriptive findings from the experiment reported in the methodology chapter. Simple statistical measures such as the mean, frequency count and percentage were used to present the data. The findings are presented based on the following sets of variables: user characteristics; search topic characteristics; search process characteristics; search term characteristics; user satisfaction characteristics; and usability characteristics.

4.1. User Characteristics

User characteristics include such variables as user status and age, subject specialism, user computer experience, thesaurus experience and user intention.

4.1.1 User Status and Age

A total of 30 users participated in the experiment, 15 academic staff and 15 postgraduate or doctoral students as shown in Table 4.1.

	Age			Total
	Less than 30	**30 - 40**	**40 - 50**	
Academic staff	0	13	2	15
Postgraduates	12	1	2	15
All	12	14	4	30

Table 4.1 Participants grouped by status and age

17 users, eight academic staff and nine postgraduate or PhD students were female and 13 users, seven academic staff and six postgraduates were male.

4.1.2 User's Subject Specialism

The distribution of the users according to their own assessment of their subject specialties is shown in Table 4.2.

Subject discipline	**No of users**
Veterinary medicine and surgery	20
Veterinary science	5
Biochemistry and molecular biology	3
Modelling and epidemiology	2
Total	30

Table 4.2 Distribution of users by their subject discipline

4.1.3 Use of Computers and CAB Abstracts

The users were asked to rate their experience in using computers (Table 4.3). Seven users (23%) had experience of more than 10 years, 16 users (53%) stated that they had 4 to 10 years of experience, four users (13.3) had 1 to 3 years of computer experience and three users (10%) had less than one year experience. An ANOVA test showed that academic staff had more computer experience than did the postgraduate and PhD students (P=0.005).

Computer use category	No of users
Less than 1 year 1 –3 years 4 – 10 years More than 10 years	3 4 16 7
Total	30

Table 4.3 Computer use experience

All users stated that they had used computer applications such as word processors, spreadsheets, the world wide web and e-mail. Twenty users (67%) indicated that they used online public access catalogues (OPACs) and seventeen users (57%) had experience of scientific database searching.

Users were also asked whether they were familiar with the CAB Abstracts database. Twenty-three users (77%) were aware of the database. Frequency of use of CAB Abstracts by those who had used the database is shown in Table 4.4.

Frequency of use	No of users
More than 3 times a month 1 –2 times a month Rarely Have never used	9 9 2 10
Total	30

Table 4.4 Frequency of use of the CAB Abstracts database

Users were asked to indicate which version of the CAB Abstracts database they had used. As can be seen from Table 4.5 most had used the Ovid CAB on the Web.

CAB version	No of users
Ovid CAB on the Web	18
CABDirect on the Web	1
SilverPlatter CAB on CD-ROM	1
Dialog CAB on the Web	0
SilverPlatter CAB on the Web	0
Total	20

Table 4.5 The CAB version used by users

4.1.4 User Thesaurus Experience

Participants were asked if they were aware of or had made use of the CAB thesaurus or other thesauri. Eight users stated that they were aware that the CAB Abstracts database had a thesaurus. However, only three users indicated that they had used the CAB thesaurus for searching. Users were also asked whether they had made use of any other thesaurus for searching. Three users had experience of using the MeSH thesaurus on the PubMed website. As mentioned earlier twenty-three participants were aware of the CAB abstracts database and had used it before.

4.1.5 User Intention

Users were also asked about the intended use of the information they were searching for. Table 4.6 shows the distribution of users' intention over the ninety search topics. MSc and PhD theses accounted for 37% of the search topics, research projects 36%, academic journal publishing 17% and teaching 11%.

User intention	No of topics
MSc/PhD Thesis preparation	33
Research project	32
Academic journal publishing	15
Teaching	10
Total	90

Table 4.6 Distribution of search topics by user intention

4.2 Search Topic Characteristics

Search topic characteristics were considered with respect to three main variables, namely topic complexity, topic familiarity and the type of search. List of users' search topics is provided in Appendix H.

4.2.1 Topic Complexity

One of the topic related variables studied was topic complexity. In order to measure the complexity of the topics provided by users, the number of search terms and number of Boolean operators used in each topic were assessed. For the purpose of comparative analysis the topics were classified as either simple or complex. Topics having more than three search terms or more than two operators were classified as complex and those with three search terms or less and two operators or less were categorised as simple topics.

Fifty-one search topics (57%) were defined as simple and thirty-nine topics (43%) as complex. Table 4.7 shows the number of topics and their distribution based on the number of search terms in each.

Number of initial terms per topic	Number of topics
Topics with 1 term	16
Topics with 2 terms	25
Topics with 3 terms	26
Topics with 4 terms	14
Topics with 5 terms	5
Topics with 6 terms	1
Topics with 7 terms	3
Total	90

Table 4.7 Number of initial terms per topic

4.2.2 Topic Familiarity

Users were asked to indicate whether or not they had previously performed searches on the topic. Fifty-seven topics (63%) had previously been searched for by users, while thirty-three (37%) topics had not.

Users were also asked to indicate the extent to which they were familiar with search topics on a three-point scale. The total number of topics within each familiarity level and status grouping are indicated in Table 4.8.

Status	Unfamiliar	Moderately familiar	Very familiar
Academic staff Postgraduates	5 10	27 25	13 10
Total	15	52	23

Table 4.8 Topics grouped by familiarity level status

As table 4.8 shows in general academic staff were more familiar with the topics than were postgraduates.

Users were also asked when they had carried out searches on the topics. Table 4.9 shows when in the past they have performed searches on the topics.

Previous search time	No of topics
Not previously searched 1 – 6 months ago 7 – 12 months ago More than one year ago	33 41 5 11
Total	90

Table 4.9 Previous search time

4.2.3 Search Type

Users were asked to determine whether they intended to conduct a *broad* or *specific* search. Broad searches accounted for 61%, whereas specific searched constitutes 39% of the topics.

4.3 Search Process Characteristics

Search process characteristics, which were taken into consideration in this study, include cognitive moves, physical moves, time spent per

search, Boolean operators used and screenshots viewed during the search process.

4.3.1 Cognitive Moves

The number of cognitive moves taken to complete each search was examined. Seven cognitive move types were identified as part of the search process. These move types together with the total number of moves of each type and the mean number per search are shown in Table 4.10.

Cognitive move types	Total number of moves	Mean number of moves per search
Term input	252	2.8
Browsing terms in mapping state	210	2.4
Browsing terms in thesaurus hierarchy state	55	0.6
Selection of terms	265	2.9
Combine search terms	192	2.1
Browsing retrieved titles	111	1.2
Query reformulation	104	1.1

Table 4.10 Total and mean number of cognitive moves per search

The mean number of cognitive moves taken per search over all searches was 13.

4.3.2 Physical Moves

In total ten physical move types were identified. The total and mean number of moves made per search are shown in Table 4.11. The mean number of physical moves made per search was 21.

Physical move types	Total number of moves	Mean number of moves per search
Perform search	254	2.8
Scroll up & down	671	7.5
Back & forward	39	0.5
Continue	192	2.1
Combine	331	3.7
Citation display & e-mail	284	3.2
Main search page	10	0.1
Previous & next page	71	0.8
Search history	15	0.2
Expand & contract	38	0.4

Table 4.11 Total and mean number of physical moves per search

4.3.3 Time

The amount of time spent by users to complete searches was assessed. In order to provide a more finely grained time analysis, two search stages were taken into consideration in terms of the time spent namely the *concept search* stage and the *search combination and result viewing* stage. The concept search stage refers to all user actions related to the input of a concept, browsing the terms suggested by the system, selecting term(s) and submitting the search terms to the database. The second stage consists of two main tasks namely combining different search terms using Boolean operators and viewing the retrieved results.

Search stage time	Mean time per search (minute/second)
Time per concept search	3" 15 sec
Time per combination and result viewing	3" 40 sec
Total time per search	6" 57 sec

Table 4.12 Mean time spent per search

The rationale for separating search stages into two main categories as above is because the time spent on browsing and selecting search terms

can be easily compared with the time spent on the rest of the search process. Table 4.12 shows the mean time (in minutes and seconds) for each search and each user in terms of the defined stages.

4.3.4 Boolean Operators

The number of Boolean operators used during the search process was assessed. As there were only two operators i.e. AND and OR available, the use of these two operators is reported. Based on the search states defined in the methodology chapter (section 3.7) there were three situations where users could choose these operators: the mapping state, the thesaurus hierarchy state, and the combine state.

All users made use of the AND operator while only 21 (70%) users introduced the OR operator while performing their searches. The use of operators was influenced by the mapping situation where one or more terms were suggested to the user. The AND operator was used in the combine state by the majority (70%) of users. The OR operator was utilised in either the mapping or hierarchical states as the terms appear in these two states. One of the main reasons for extensive use of operators, in particular the OR operator, lies in the fact that the Ovid search interface makes operators explicitly available at search term selection and search combination situations. Table 4.13 shows the total number of operators used in each of the search states.

Operator	Mapping State	Hierarchical thesaurus State	Combine State	Total
AND	28	4	97	129
OR	52	11	0	63
Total	80	15	97	192

Table 4.13 Number of times operators were used in each state

As can be seen from Table 4.13 users made use of the OR operator mostly in the mapping state where the user is provided with a list of terms for browsing and selection.

Table 4.14 shows the mean number of operators used per search and per user. On average a total of two operators were used in each search.

Operator	Mean per search
And	1.4
Or	0.7

Table 4.14 Mean number of operators per search

4.3.5 Screenshots Viewed

The number of screenshots users viewed while interacting with the search interface was assessed. Screenshots relating to search concepts, search term combination and result viewing were counted. The mean number of screenshots per search for these stages is shown in table 4.15.

Screenshots	Mean per search
Input, browsing and selecting terms	9.2
Combining terms and viewing results	7.7
Total search screenshots	16.8

Table 4.15 Mean number of screen shots viewed per search

On average users viewed 9 screenshots while browsing and selecting search terms and 8 screenshots viewing search results and combining terms. The table indicates that on average just over half of the screen shots viewed were associated with the input, browsing and selection of search terms.

4.4 Search Term Characteristics

Search term characteristics encompass the analysis of all search terms browsed and selected, query reformulation and expansion, and the characteristics of the ways in which users' terms mapped to the thesaurus. List of users' initial search terms is provided in Appendix I.

4.4.1 Initial, Browsed and Selected Search Terms

All initial search terms entered by users were counted and the results are shown in Table 4.16.

Number of initial search terms	252
Mean initial terms per search	2.8

Table 4.16 Total and mean number of initial search terms

On average users entered around three search terms when initially formulating their queries. Figure 4.1 shows the number of initial terms for all searches.

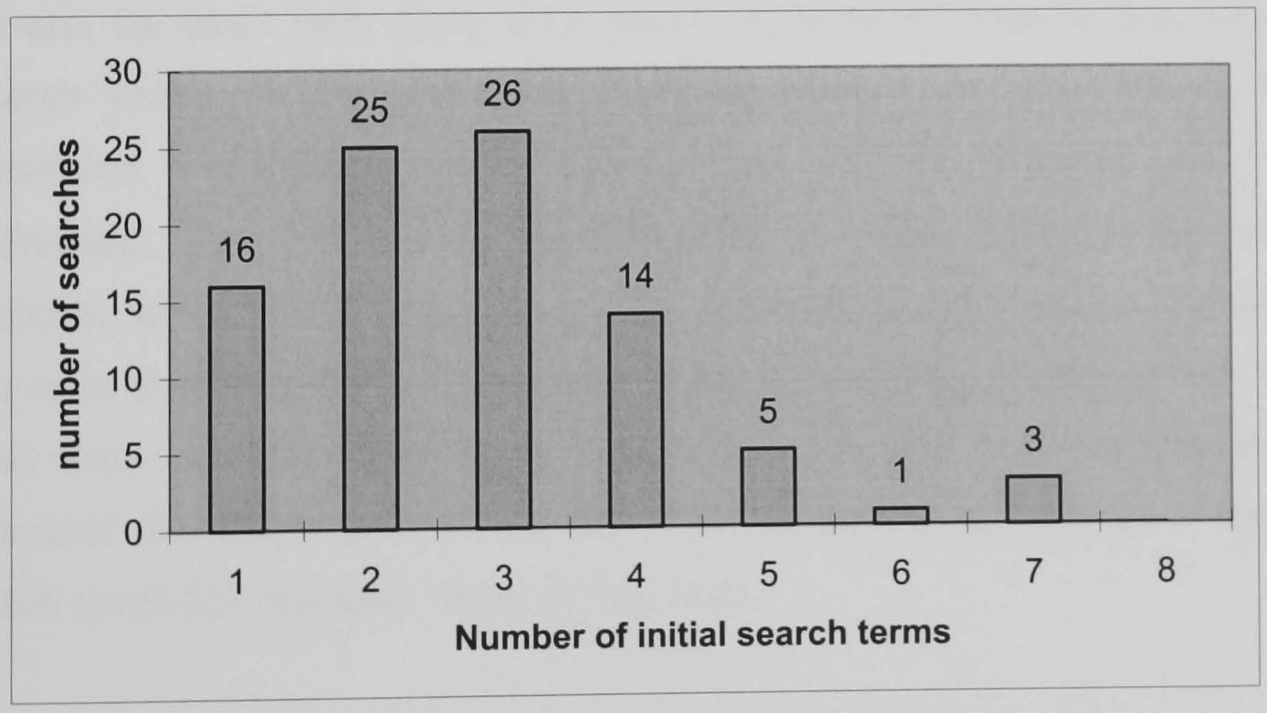


Figure 4.1 Distribution of searches by initial search terms

Search terms browsed and selected by users were also assessed to evaluate the extent to which each search state contributed to the search term selection process. There were two states at which users could

browse and select search terms depending on their search topic and the mapping facility. These two stages were the thesaurus mapping state and the thesaurus hierarchical state.

Table 4.17 provides details of the browsed and selected terms within these states. While users entered the thesaurus mapping state for all searches, only for 36 searches did users decide to access the thesaurus hierarchical state.

Search state	Terms browsed	Terms selected	Mean per search	
			Browsed	Selected
Thesaurus mapping state	2768	396	31	4.4
Thesaurus hierarchical state	637	87	18	2.4

Table 4.17 Total number of terms browsed and selected
(These numbers were based on only the 36 searches for which this state was entered)

As can be seen from table 4.17 users preferred to browse and select search terms in the thesaurus mapping state than they did in the thesaurus hierarchical state. There may be two reasons for this behaviour. The first is the fact that 64% of users' initial terms were matched exactly with thesaurus terms, therefore users decided to select the term then and there rather than going into details of the term in the next state. Another explanation for this could be that the mapping state provided the users with a rich set of thesaurus terms and as a result users could find relevant terms at that state.

Figure 4.2 shows the number of browsed terms for all searches. As can be seen over 66% of searches fall within the range of 11 to 50 browsed terms.

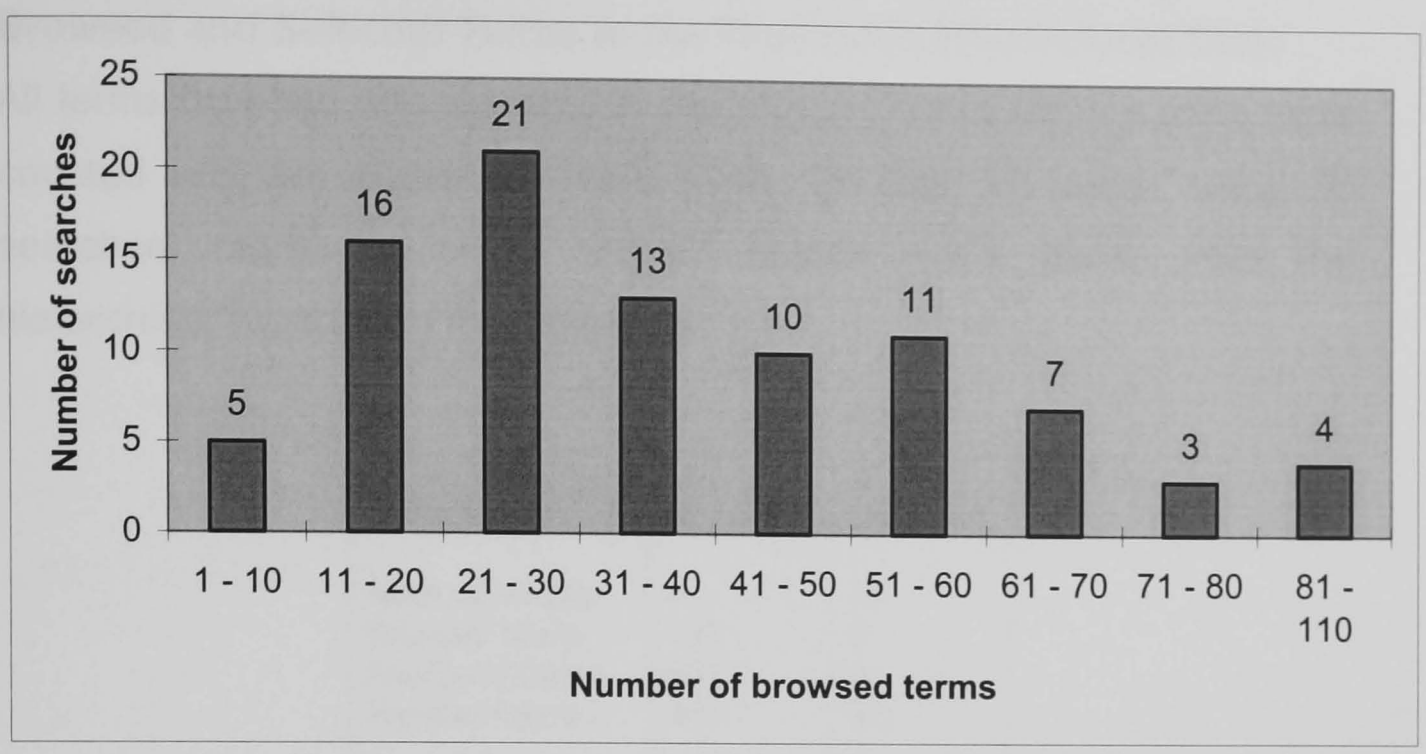


Figure 4.2 Distribution of searches by browsed terms

Figure 4.3 shows the frequency distribution for the number of selected terms for all searches. Users selected between two to seven search terms in over 77% of the searches.

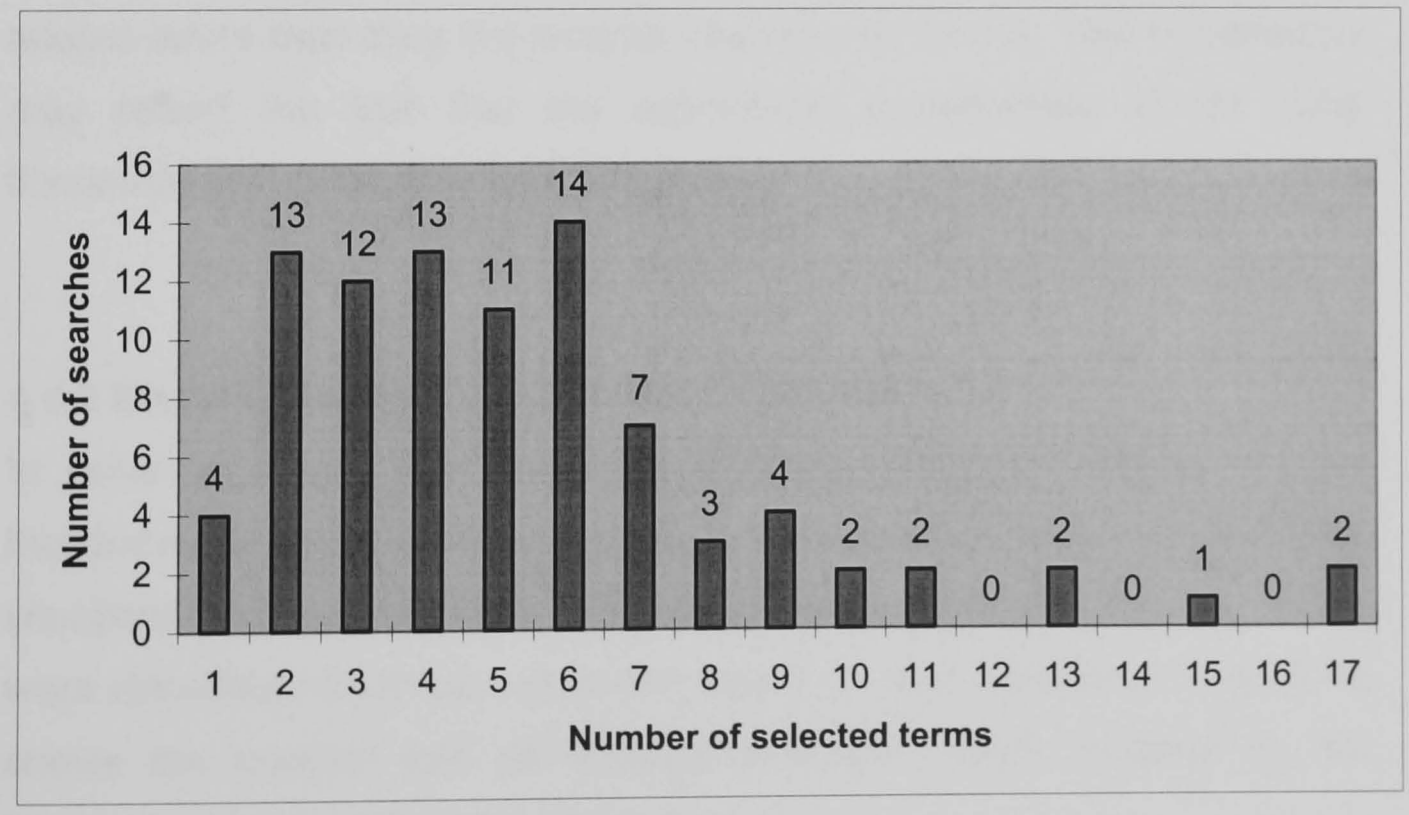


Figure 4.3 Distribution of searches by number of selected terms

Browsed and Selected Terms in the Thesaurus Hierarchical State

All terms browsed and selected in the thesaurus hierarchical state were counted and are shown in Table 4.18. In total 20 users, using 36 searches, chose to either browse and/or select terms from the hierarchical structure of the thesaurus.

Term types	Browsed	Selected
Main descriptor	--	51
Broader terms	138	6
Narrower terms	237	8
Related terms	262	22
Total	637	87

Table 4.18 Browsed and selected terms in the thesaurus hierarchical state

It should be noted that users who browsed the hierarchical structure of the thesaurus in most cases found the top term or the main descriptor as a good search term. It is also interesting to note that users selected more related terms than they did broader or narrower Terms. This observation may reflect the fact that the associative relationships in the CAB thesaurus are particularly well-developed.

4.4.2 Mapping Users' Terms to the Thesaurus

In order to assess the extent to which users' terms mapped to the thesaurus, all terms entered by users were analysed and their mapping situations are reported here. All search terms entered initially by users were classified into three categories based on their mapping. Table 4.19 shows the number and percentage of search terms mapped to the thesaurus based on these match types.

Match type	No of terms	%
Exact match	160	64
Partial match	45	18
Statistical match	47	18
Total	252	100

Table 4.19 Number and percentage of user search terms matched to the thesaurus

A brief description of different match types, together with some examples is provided in the following.

Exact matches: While the notion of 'exact match' may appear to be self explanatory, it should be noted that within this study those terms with a scientific name or name of a particular family of animals have been treated as exact matches. For example:

User term: **bovine, equine**
Thesaurus term: **bovidae, equidae**

Singular noun terms which matched to the plural form in the thesaurus were also treated as an exact match. For example:

User term: **mouse, animal**
Thesaurus term: **mice, animals**

The reason for these decisions lies in the fact that the Ovid search interface automatically refers users to the plural form.

Matching acronyms were also considered as being exact matches. For example:

User term: **FMD**
Thesaurus term: **Foot and Mouth Disease**

Partial matches: These types of match were typically between one query word and a multi-word thesaurus descriptor. Partial matches also included user terms which have part-of-speech difference from thesaurus descriptors. This entails matched terms with different endings or terms that are adjectives of the initial terms or have been subject to a stemming process. Examples:

User term: **porcine**
Thesaurus term: **porcine adenovirus**

User term: **pelvic**, **behavioural**
Thesaurus term: **pelvis**, **behaviour**

Statistical matches: These include thesaurus descriptors which are matched based on a statistical analysis to determine which descriptors tend to occur in documents containing the query term. When the system is unable to find an exact or partial match to the user term, it suggests thesaurus terms which co-occur frequently with the user term. For example:

User term: **hyperalgesia**
Thesaurus term: **pain**

User term: **Dolly**
Thesaurus term: **sheep**, **cloning**

4.4.3 Free Text Search

In all three mapping situations there was an option called *search as keyword* whereby users could search their query terms as free text rather than thesaurus search. They could choose this option together with exact, partial or statistical match to enlarge their result sets. Table 4.20 shows the match types together with the number of times the *search as keyword* option was used.

Match type	No of search terms
Exact match Partial Match Statistical match	31 27 32
Total	90

Table 4.20 Number of times the *search as keyword* option was used

An analysis of the use of this option indicated that users utilised this feature in 55 (61%) searches, involving a total of ninety search terms. The 'search as keyword' option appears to be used fairly equally across the matching states. This differs from thesaurus term matching where exact match is clearly the predominant category. There was a difference observed in the way users made use of free text searching in different matching situations. In the exact and partial match cases they selected the free text option as the secondary source of additional terms which would improve their search while in the statistical match situation they relied more on free text searching and chose this option as a primary source for their search.

4.4.4 Query Reformulation

Query reformulation encompasses any effort to modify or expand the query. To provide a detailed view of query reformulation instances, query modification and query expansion are discussed separately in the following.

Query Modification

Query modification is defined as a process within which users delete and/or modify initial search terms. In total 21 users reformulated 36

queries. The reasons for query modification were elicited during their search and the results are shown in Table 4.21.

Modification reason	Number of queries modified
No hits on initial query Too few hits on initial query Too many hits on initial query Unsatisfied	11 12 7 6
Total	36

Table 4.21 Distribution of modified queries by modification reason

The modified queries were also analysed in terms of the stage at which users were motivated to reformulate their queries. As is shown in Table 4.21 two third of modified queries were associated with insufficient hits. In 25 of the reformulated queries users examined the retrieved titles first. Of the 36 reformulated queries, 6 queries were reformulated by language limit and 1 query with date limit.

Strategies users adopted to modify their queries were also analysed. The analysis was based on the notion of search term tactics proposed by Bates (1979). The three main tactics which constituted the basis for the analysis of modified queries are:

- Super: to move upward hierarchically to a broader term
- Sub: to move downward hierarchically to a more specific term
- Relate: to move sideways to a coordinate term

Table 4.22 shows the distribution of reformulated queries by reformulation strategy. Note that seven reformulated queries were not accounted in this table as they were reformulated by date or language limits. The number of the *Super*, *Relate* and *Sub* strategies appears to be

roughly proportional to the main reasons provided by users for query modification (i.e. insufficient hits or too many hits).

Status	**Query modification strategy**		
	Super	Sub	Relate
Academic staff	7	5	4
Postgraduates	7	2	4
Total	14	7	8

Table 4.22 Distribution of reformulated queries by reformulation strategy and status

As can be seen from Table 4.22 academic staff and postgraduates adopted the same number of *Super* and *Relate* strategies while academic staff made more use of *Sub* strategy than did postgraduates.

One of the issues of interest in query modification relates to the source of search terms. The modified queries were analysed based on their source and three main sources were identified namely: the thesaurus, the users' own terms, and the terms selected from the result set. Table 4.23 shows the number of terms selected during all modifications totalled for each term source.

Modification term source	**No of terms**	**%**
Thesaurus	18	47
Users' own terms	13	35
Terms selected from retrieved items	7	18
Total	38	

Table 4.23 Sources of search terms in the reformulation process

The above table indicates that the thesaurus was the first source of search term selection for modifying queries followed by users' own terms.

Query Expansion

Query expansion is a process whereby users supplement their initial search terms with additional terms in order to improve the retrieval performance. The total number of query expansion instances and total number of expansion terms was analysed. In total users expanded 68 (76%) searches out of ninety. Table 4.24 provides the total and mean number of initial terms in expanded and unexpanded searches. The table shows that expanded searches had more initial terms than did the unexpanded searches.

Type of initial search terms	No of terms	Mean per search
Initial terms (68 expanded searches)	196	2.8
Initial terms (22 unexpanded searches)	56	2.5

Table 4.24 Total and mean number of initial terms in expanded and unexpanded searches

One explanation for the finding that 76% of searches were expanded could be the fact that in a thesaurus-enhanced search environment, where users are provided with terms semantically related to their initial search terms, they tend to select and add more terms to their queries. In order to assess the extent to which the thesaurus provided users with new terms, a comparison of users' initial terms and the thesaurus terms selected by users for query expansion was carried out. It was shown that the number of new terms selected for expanded searches varied from 1 to 13, with a mean of 2.8. An evaluation of expanded terms at the user level demonstrates that users on average selected 6 new terms while expanding between 1 to 3 searches. As Figure 4.4 shows in the majority of expanded searches users selected between 1 to 6 query expansion terms.

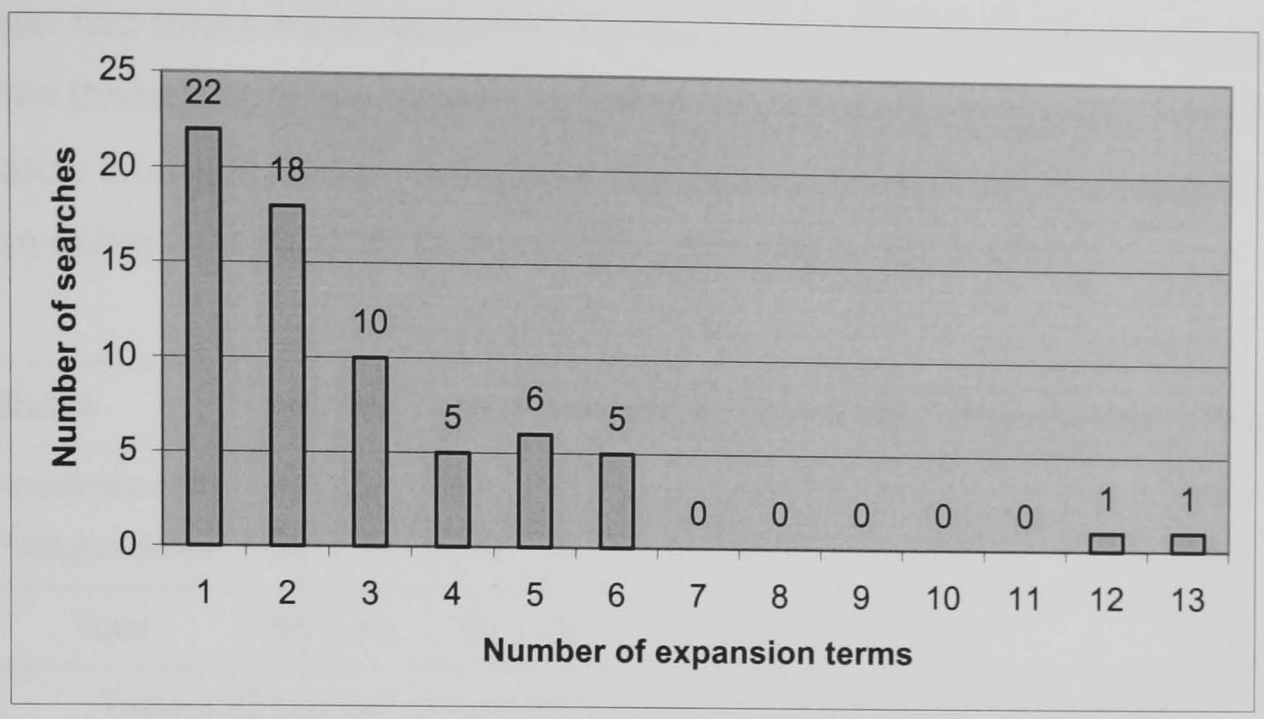


Figure 4.4 Distribution of searches by the number of query expansion terms

To evaluate the extent to which the thesaurus contributed to the query expansion process, it is important to investigate query expansion terms selected from the thesaurus and their relationship to users' initial search terms. Of 193 expansion terms selected by users, 157 query expansion terms were selected at the thesaurus mapping state and the remaining 36 terms were selected at the thesaurus hierarchical stare. Since thesaurus notations, namely BT, NT, and RT, are not specified at the thesaurus mapping state, query expansion terms which were selected at this state had to be individually examined in the context of the tree hierarchy to establish their relationship with user terms. Table 4.25 shows the six types of terms identified in analysing users' query expansion terms in their hierarchical context. Two types of these terms, namely statistical and alphabetical, require clarification. Statistical terms include those thesaurus terms that did not have any thesaurus-like relationships to users' terms but were judged by the system to be relevant to the users' search based on co-occurrence analysis terms. The relationships between statistical match terms to users' terms cannot be established, as they have been suggested as the result of statistical co-occurrence analysis of terms. For instance, there is no way of establishing a relationship between *Ecoli 0158* and *China*. But in the database records

these two terms have frequently co-occurred. Alphabetical terms include those thesaurus terms located in the thesaurus alphabetical list. The Ovid search interface to the thesaurus hierarchical relationships displays the term in the context of its alphabetical order at the same time.

Status	NT	RT	BT	Statistical	Synonyms	Alphabetical	Total
Academic staff	33	24	8	25	12	2	104
Postgraduates	25	30	11	14	7	2	89
Total	58	54	19	39	19	4	193

Table 4.25 Number of expansion terms by relationship type and status

Table 4.25 shows that narrower and related terms taken together constitute around 58% of the query expansion terms selected by users, reflecting the fact that narrower and related terms appear to represent the largest category of query expansion terms. Statistical relationship terms accounted for around 20% of expansion terms. It is interesting to note that while the terms identified as being statistical are thesaurus terms, they do not represent any thesaurus-like relationships to the users' terms. Around 10% of query expansion terms were broader terms and synonymous terms accounted for 10% of the total query expansion terms. Table 4.25 also shows that academic staff selected more narrower, statistical and synonymous terms than did postgraduates. The postgraduates, however, selected more broader and related terms than did academic staff. The differences in the number and type of terms selected by both groups point to different patterns of search behaviour.

In order to evaluate the search state in which users were more likely to select expansion terms, all query expansion cases were analysed. The results are shown in Table 4.26.

Expansion State	Status		Number of Searches
	Academic staff	Postgraduate	
Thesaurus mapping state Thesaurus hierarchical state	20 15	21 12	41 27
Total	35	33	68

Table 4.26 Number of query expansion searches at different states

The above table suggests that users in 60% of cases preferred to select their expansion terms from the thesaurus mapping state while choosing to select their expansion terms from the hierarchical state in just under 40% of cases.

While the table shows that in general academic staff members entered the thesaurus hierarchical state more than did postgraduates, this difference was not statistically significant. It should be noted that the choice of browsing and selecting terms from either the mapping or hierarchical state was dependent on the mapping situation. Most users browsed the hierarchy when their search terms had an exact match in the thesaurus. This was due to the fact that there was only one term the exact match term and they tended to enter the hierarchy state. Since partial or statistical matches immediately brought up a list of relevant terms, users often made their term selection at the thesaurus mapping status rather than going into the hierarchical state.

The 22 unexpanded searches were also analysed to see if there were any common characteristics among them. A detailed analysis of the unexpanded queries indicated that they tended to belong to one of the two types:

- A number of the unexpanded queries contained terms which matched exactly with thesaurus terms, so there was little need to expand the queries. 14 out of the 22 unexpanded queries fell into this category.

- There were also 8 queries which consisted of a single word and for which no exact or partial thesaurus match was found. In these cases the user decided to search using the free text search option.

Another observation in relation to unexpanded searches was that the mean number of initial terms in these queries was 2.5 pointing to the fact that unexpanded queries had less initial search terms than expanded searches. However, this difference was not found to be statistically significant.

4.5 Search Result and Search Term Satisfaction

4.5.1 Search Results Relevance

The total and mean number of records retrieved, browsed and selected per search was assessed and the results are shown in table 4.27. On average for each search users retrieved 75 records, browsed 16 record titles and selected 55 records as being relevant for sending to their e-mail addresses. Only 11 users chose to view some abstracts of retrieved sets.

	Total number of records	**Mean per search**
Total records retrieved	6750	75
Record titles browsed	1414	16
Selected and e-mailed records	4977	55

Table 4.27 Total and mean number or records retrieved, browsed and selected per search

As can be seen from this table, on average users sent 73% of the retrieved records to their e-mail address as being thought relevant.

In addition, users judged the relevance of the retrieved set based on a three-point scale. This is the most common relevance assessment scale reported in recent research. Table 4.28 quantifies the results for all the topic searches.

Relevance rating	No of searches
Very relevant Partially relevant Not relevant	60 29 1
Total	90

Table 4.28 Users' relevance rating

In order to evaluate users' perception on the relevance of results sets from a different perspective, users were also asked if the results they viewed generally matched the topic they were interested in using a three-point scale. Table 4.29 shows the users' perceptions towards the match of results with their search topics.

Match of results to topics	No of topics
Very good match Fairly good match Poor match	49 37 4
Total	90

Table 4.29 Match of results to users' topics

A comparison of tables 4.28 and 4.29 shows a slight difference in the way users judged the result sets. They judged sixty searches as being very relevant, they commented that only forty-nine searches provided very good match to their topic. This finding indicates that topics are not the only bibliographic surrogates on which users base their relevance judgement. It also points to the fact that while the retrieved results may not exactly match topics, they can still be very relevant to the user's information need.

4.5.2 Search Result Satisfaction

User ratings of how satisfied they were with the number of search results were assessed using a four-point scale. Table 4.30 shows satisfaction

ratings for all searches. A four-point scale was used to clearly delineate the distinct levels of user satisfaction and provided appropriate data for non-parametric statistical analysis.

Result satisfaction	No of searches
Very satisfied	48
Fairly satisfied	35
Unsatisfied	6
Very unsatisfied	1
Total	90

Table 4.30 Users' ratings of result satisfaction in terms of the number of results

As Table 4.30 shows just over half the searches were rated as *very satisfied* with 39% rated as *fairly satisfied*. In only 8% of searches did users state that they were unsatisfied or very unsatisfied.

From the gender perspective we note an interesting trend in users' satisfaction with the number of search results. Table 4.31 presents the satisfaction ratings by gender across all searches. As can be seen from Table 4.31 female users were in general more satisfied with search results than were male users.

Gender	Very satisfied	Fairly satisfied	Unsatisfied	Very unsatisfied	Total
Female	36	12	3	0	51
Male	12	23	3	1	39
Total	48	35	6	1	90

Table 4.31 Result satisfaction rating and gender

4.5.3 Search Term Satisfaction

Users were asked to comment as to whether the thesaurus terms they selected were close to their original search topic and how satisfied they were with the number of search terms. The findings relating to term closeness and term satisfaction are reported in the first section. Then, an

evaluation of users' perception of search terms they selected from the thesaurus was reported based on the following questions: a) whether or not the thesaurus suggested additional search terms, b) whether or not the additional thesaurus terms were useful, and c) whether or not the users had been aware of the additional terms suggested by the thesaurus at the beginning of the search.

Term Closeness and Term Satisfaction

The users' perception of the closeness of search terms to their original search topics was measured on a three-point scale: very close, fairly close, not at all. Table 4.33 shows the users' ratings of term closeness.

Closeness level	Number of searches
Very close Fairly close Not at all	47 38 5
Total	90

Table 4.33 Users' rating of term closeness to original search topics

The users found that for just over half the searches the terms provided were *very close* to their topic, and in a further 42% of searches *fairly close* and only 6% of searches were mentioned as being not at all close.

Users' satisfaction with the number of search terms they selected from the thesaurus was evaluated using a four-point scale: very satisfied, fairly satisfied, unsatisfied, very unsatisfied. The reason for choosing a four-point scale is the fact that the data can be easily used for the purpose of statistical analysis. Table 4.34 summarised users' response.

Satisfaction with the number of terms	No of searches
Very satisfied Fairly satisfied Unsatisfied Very unsatisfied	41 41 7 1
Total	90

Table 4.34 Users' satisfaction with the number of terms selected

Users' satisfaction rating with respect to gender is shown in table 4.35. As can be seen, female users were generally more satisfied with the number of search terms they selected.

Gender	Very satisfied	Fairly satisfied	Unsatisfied	Very unsatisfied	Total
Female Male	30 11	18 23	3 4	0 1	51 39
Total	41	41	7	1	90

Table 4.35. Users' term satisfaction and gender

Additional Search Terms

Of the 90 searches, the thesaurus provided additional terms for 74 (82%) searches while no additional terms were found in 16 searches. All thirty users found additional terms from the thesaurus for at least one of their search topics.

Users were asked to comment on the usefulness of the additional search terms. Of the 74 searches for which the thesaurus provided additional terms, 65 searches were judged by users as having provided useful additional terms. For 9 searches the users did not find the additional terms to be useful. In relation to users' awareness of the additional thesaurus terms at the onset of the search, in 37 searches users mentioned that they had not been aware of the terms at the beginning. Of these 37 searches, 23 belonged to postgraduates while 14 searches

belonged to academic staff. The users were also asked about the reasons for which they chose additional terms, with 27 users providing comments for sixty searches. The answers to this question were categorised broadly into five main responses and are shown in Table 4.32.

General categories of comments	**Number of searches**	**Examples of users' comments**
Narrow down the search	28	• It narrowed general search to very specific aspect • To narrow down the field I was looking at • It makes the search more specific for my field of interest • Narrower search achieved • To limit the number of records • To restrict dataset • More specific terms • It narrowed down my search through suggesting a new angle
Broaden the search	13	• It provided broader search terms • Better and broader definition of term wanted • To broaden selected citations • Spread the range of topics • To return a larger number of search results • Ensure complete search
Term relatedness	11	• Seemed relevant to the information that I was hoping to obtain • Relevant, I wish I had thought of them previously • They were relevant and alternative verb or adjective forms of my search terms • To obtain relevant references • Seemed relevant
Initial search result	3	• Poor response to first attempted search • Term used had no matches or very limited number of matches • Confirm that the original terms were correct • Initial search unproductive
Miscellaneous	5	• Better terms • Exact terms suggested • They matched original terms • More accuracy, richer vocabulary when mine was limited • Additional terms were not selected but were useful to confirm that the original terms were correct

Table 4.32 Reasons for choosing additional thesaurus terms given by users

As can be seen from the above table the three main reasons users indicated that they chose additional thesaurus terms were *narrowing down* (47%), *broadening* (22%) and *term relatedness* (18%). This finding demonstrates that end-users understand thesaurus relationships and the ways in which they can be used to broaden or narrow down their searches.

4.6 Thesaurus and Search Interface Usability Characteristics

The post search interview elicited information about the usability of the thesaurus and of the search interface as well as of the users' general experience. Since users were interviewed following the completion of all three searches, the results reported here are based on comments made by 30 individual users.

4.6.1 Thesaurus Usability

The participants' comments on the ease of thesaurus browsing and navigation were recorded using a five-point scale that ranged from 1 to 5 on which 1 indicated *very difficult* and 5 indicated *very easy*. The usability scales reported in the literature range from 3 to 7 and in this study an average of these scales was adopted. The users' rating is shown in Table 4.36.

Ease of browsing and navigation	No of users
Very easy	17
Easy	11
Fairly easy	2
Difficult	0
Very difficult	0
Total	30

Table 4.36 Ease of thesaurus browsing and navigation

The mean rating for the ease of thesaurus browsing and navigation was 4.5 with the majority of users (93%) assessed the thesaurus browsing and navigation as either very easy or easy. No one felt that thesaurus browsing/navigation was difficult.

In order to assess other major aspects of thesaurus usability the following questions were asked during the interview session:

1. What, if any, effect did browsing the thesaurus have on your selection of search terms?
2. What kinds of problems and difficulties did you experience during thesaurus browsing?
3. What type of help do you think you needed while browsing the thesaurus?

The responses to each of the above questions were content analysed and the core themes and categories were defined. The order of the results is based on the sequence of the above questions.

What, if any, effect did browsing the thesaurus have on your selection of search terms?

All academic staff and postgraduates stated that thesaurus browsing had some effect on their search term selection. In order to describe these effects, responses to the above question were analysed and categorised according to participants' status, namely academic staff or postgraduates. Table 4.37 summarises the results.

Academic staff	Postgraduate students
Narrow down • Provision of narrower search terms • Make the search more specific • Narrow down the search • Additional ways of making a search specific • Provision of a broad range of terms for narrowing down my search **Narrow down/broaden** • Depending on the search topic it can help to search for a narrower field or more general field • It narrowed down and in some searches widened; it depends on what you start with • Broad and narrower terms depends on topic **Familiar/Unfamiliar topics** • It is useful when I am less familiar with a topic • It would be useful for topics which I am not very familiar with • It provides an accurate vocabulary I would not know the right term, It helped me to focus, it directed me **Miscellaneous** • Made the search quicker as it provided the other terms that I intended to input • It made the search easier as you can select terms once as compared to individual • A lot more options to choose from	**Narrow down** • Narrow down and make the search more specific • Helped me make my search more specific • Make the search more specific to field of my interest **Broaden** • To broaden the search, inclusion of more search terms • It widened the search • To perform a broader search • Broadened the search • It broadened the term I used **Narrow down/broaden** • Depending on the search topic it can improve the search through broadening and narrowing • Depends on the original search terms to broaden or narrow down **Term awareness** • Coming up with terms which I was not aware of because I was not familiar with the area • Coming with terms that you might not have thought yourself • Provision of terms that I would not have thought of using /provision of additional terms **Miscellaneous** • Provision of additional terms • More relevant terms than my initial terms • Introducing related areas for search, useful • Wider and better way to search and better terms

Table 4.37 Users' comments on thesaurus browsing effects

Both groups commented on a number of common thesaurus browsing effects, such as broadening or narrowing down the search, provision of additional search terms, etc., as listed in Table 4.37. However, there were certain areas in which the focus of the groups' comments differed. For instance a number of academic staff stressed the fact that whether the thesaurus was used for broadening or narrowing down the search depends largely on the topic the searcher starts with. Some staff also commented that the thesaurus would be more useful to those who are less familiar with the search topic at hand. This comment can be compared to what some postgraduate participants stated about their lack of awareness of thesaurus terms at the beginning of the search.

An analysis of the comments made by both groups shows that the thesaurus provided different functions for academic staff and postgraduate students. Based on the comments and also data gathered on users' topic familiarity in the pre-search questionnaire the following statements can be made: Postgraduate subjects were less familiar with topics, tended to comment more on the *broadening* effect of the thesaurus, found the thesaurus browsing as useful in providing terms they were not aware of at the onset of the search, and appreciated the thesaurus for introducing them to related terms and areas. Academic staff were more familiar with their search topics, tended to comment more on the *narrowing down* effect of the thesaurus, and the usefulness of the thesaurus in providing alternative terms.

What kinds of problems or difficulties did you experience during thesaurus browsing?

Six academic staff and 3 postgraduate students noted that they did not experience any difficulties. The comments made by the rest are categorised and listed in table 4.38.

As can be seen from Table 4.38 in general postgraduate participants experienced more problems and difficulties in terms of thesaurus browsing, use of system features and their uncertainty about thesaurus terms. An explanation for this could be the finding reported in section 4.1.3 that academic staff had more computer experience than did postgraduate students.

Academic staff	Postgraduate students
Problems associated with narrowing down • Selection of too narrow terms • Too specific terms exclude a large number of references **Uncertainty about broader or narrower terms** • You do not know whether you will be provided with broad or narrow terms • Unsure as to which term to select or which term would be the best • In the thesaurus hierarchy I am not sure which terms narrow down the search and which terms expand, I know some indention had an implication for narrow **Interface** • Navigating around the buttons: the their size and position • Number of hits in front of each term a bit confusing	**Problems associated with broadening** • My terms were too general • Sometimes there are too many terms for a certain situation • The large number of results **Problems associated with narrowing down** • Sometimes confusing as you don't know you are narrowing your search too much **Uncertainty about broader or narrower terms** • Struggling to identify what each term in the thesaurus actually meant and how to combine them • Sometimes it brought up terms I did not think as applicable to my research • A slight ambiguity to differentiate between thesaurus term and the keyword option • Lack of clear understanding as to what I will see when I select terms **Use of operators** • A little bit difficulty in using AND and OR • Using AND sometimes you don't get the results you expect

Table 4.38 Users' problems and difficulties during thesaurus browsing

What type of help do you think you needed while browsing the thesaurus?

Ten academic staff and six postgraduate students stated that that they did not need any help while browsing the thesaurus. The comments made by the rest of participants are summarised in Table 4.39.

Academic staff	Postgraduate students
• A reminder that the searcher can change AND and OR. • I may need a tutorial • During the selection of thesaurus terms specially when you have broader and narrower terms close to each other • A guide as to which terms broaden or narrow your search	• A guide as to what the Explode option does • How terms are going to help you • Distinguish between the suggested terms which relate to the searcher's initial term and those terms with have no direct relevance • An initial guide as to what everything means and how to go about it • Unsure whether I should click on the term or tick the box next to the term • Differentiating between AND and OR

Table 4.39 Users' suggestions for help features during thesaurus browsing

A number of users from both groups pointed out the need to provide guidance as to how to use Boolean operators. They also mentioned that they needed some indication on the nature of broader and narrower terms and the influence they may exert on search results.

4.6.2 Interface Usability

The usability of the interface was assessed based on three questions, two of which were on a five-point scale. The last question of this section was descriptive to elicit users general perceptions.

The first question related to ease of learning and use of the interface. Table 4.40 summarises the extent to which users thought the interface was usable and learnable.

Ease of learning and use	No of users
Very easy	10
Easy	19
Fairly easy	1
Difficult	0
Very difficult	0
Total	30

Table 4.40 Ease of learning and use of the interface

The mean rating was 4.3 with all but one user commenting that the search interface was either very easy or easy to learn and use.

The second interface usability question was associated with the ease of using commands within the interface. Table 4.41 shows the users' rating.

Ease of command use	No of users
Very easy	17
Easy	13
Fairly easy	0
Difficult	0
Very difficult	0

Table 4.41 Ease of command use within the interface

The mean rating for ease of command use was 4.6 and all users evaluated the use of commands as being either very easy or easy.

A simple comparison of the above two tables shows that learning and using the interface and search commands available within the interface was an easy task. Users also described their experience of using search commands in the interface. Sixteen users did not provide any comments on the use of commands. Table 4.42 attempts to summarise the rest of the users' account of their command use experience.

Academic staff	Postgraduate students
• They are easy as long as you know what kinds of command are there • The *continue* button is at the top, the user has to scroll while it should also be available at the bottom • The *continue* button should also be available at the right side • *Continue* button was not easy to use • Buttons are self-explanatory • It is a nice interface, I am frustrated by the lack of the NOT function • Buttons i.e. *continue, combine,* and *perform search* need to be slightly bigger • Very straightforward	• For the *continue* button you have to scroll down • If you know what each bit does, it is straightforward • Self-explanatory • Not sure how to use AND and OR • *Continue* button should be easily available and users do not have to change their status to get access to that • *Continue* button should be bigger • It would be useful to have bar at the top for showing how to get access to citations rather than scrolling down • You have to get to the mouse to click on the perform search command rather than simply pressing the 'enter'

Table 4.42 Users' perception of the search commands

Although users evaluated the interface as easy to use and learn, from the comments above it can be observed that users experienced some level of challenge. The main usability problems pointed out by users were: the size and position of buttons, in particular the *continue* button which is one of the most frequently used commands on the Ovid search interface, Boolean operators and their usage, and access to the results page.

The following question intended to elicit users' general feelings about the interface and its features:

What did you like or dislike about the interface?

Table 4.43 summarises the likes and dislikes mentioned by the two user groups.

Academic staff	Postgraduate students
Likes • It is easy to use and self-explanatory • You can choose several different combinations of search terms without losing your original ones • It is not overly fancy • E-mail facility • The format and the background colour of the interface • It is not cluttered • Presentation of references • Fonts are easy to read and follow • Presentation of search terms **Dislikes** • I do not like to search terms separately • *Continue* and many of the buttons do not tell you had pressed them • Too much scroll up and down • Small size of the buttons • The lay-out was not intuitive • Lack of help as to which terms expand or narrow the search	**Likes** • Easy to use/user friendly • Presentation of related and specific terms • The combine option is excellent • Quick access to results • Checking the American spelling • Well-laid out • It is uncluttered **Dislikes** • Having to enter term separately • Combine options are better be ordered by words rather than numbers • The lay-out is not user-friendly • It is not intuitive • The button labels are not self-explanatory • AND option was confusing • Some buttons did not show if the system was functioning

Table 4.43 Interface likes and dislikes

When comparing the comments made by both groups, a number of common interface features can be observed which provide suggestions for improvement to the interface. Some of the main points were the size of and labels on buttons, reducing the scrolling in particular during thesaurus browsing, help option for thesaurus terms and their usage, and more description and guidance on the use of Boolean operators while selecting thesaurus terms.

Do you have any other comments about the thesaurus, interface or the search in general?

Table 4.44 summarises comments made by both groups of users.

Academic staff	Postgraduate students
Positive comments Thesaurus • The thesaurus allows more accurate hit rate • The thesaurus provides a more systematic approach in searching • The thesaurus has much more variety in terms of covering animal species • It does seem to come up with relevant terms in terms of cross-referencing the terms you put in • It can narrow down the search quite quickly • I am enormously impressed by the number of terms shown and the number of references • The thesaurus allows me to change my mind and search for different subjects • It was useful as it is difficult to find a specific term; you enter a term and it comes up with related and synonymous terms that you have not thought of before • It is difficult and time consuming to narrow down a search and the thesaurus can help to speed up the process • The thesaurus is helpful • I should use the thesaurus more now Interface • The facilities of the database were fine • The interface is easy to use • The combination option was very helpful • It was better than any other databases I have used **Criticisms** • The number of results was very limited • Long list of terms and the one I was looking for was the last term at the bottom	**Positive comments** Thesaurus • It may sometimes remind you of the term you were not thinking before • It is quite easy to use • Thesaurus is a sort of shortcut that provides you with terms that you might want to modify at later stages • It makes the search more specific • I like the way thesaurus shows terms related to each other • It does help really • Much easier than I thought, I would use it • It gives you ideas • It is definitely useful • It was beneficial • The thesaurus provides a context for a particular term; I searched for resistance and it came up with different types of resistance in different contexts • It provides a wider choice of terms • Search can be done more efficiently using thesaurus, • The thesaurus helped more to broaden my specific target to get some results **Criticisms** • System crash • The thesaurus is not going to be of help for very specific searches

Table 4.44 Users' general comments on the thesaurus, interface and the search

An analysis of the comments in Table 4.44 provides an insight into thesauri usability and functionality from the end-user perspective. The functions of the thesaurus based on users' opinions can be generally described as follows:

- Introducing new ideas and aspects related to a particular search

- Provision of contextual information for a search term
- Provision of alternative ways of formulating a query
- Useful in terms of broadening or narrowing down a search
- Supporting ways of looking at a search from different angles
- Usefulness of semantic relationships in the thesaurus for guiding users
- Reminding users of terms they were not thinking of at the beginning of a search

Most users commented that they have not previously been introduced to the thesaurus as a searching tool. They also mentioned that they would use it for future searches.

While most users made positive comments about the thesaurus, the way thesaurus terms are displayed and other usability aspects such as those reported in the previous section should be taken into consideration in order to present end-users with a more usable and functional interface.

4.7 Summary

The aim of this chapter was to present the descriptive and qualitative findings obtained from the experiment reported in Chapter 3. Characteristics of users, search topics, the search process, search terms and usability evaluation have been presented. The next chapter provides the results of statistical data analysis and hypotheses testing.

Chapter 5: Statistical Data Analysis and Hypothesis Testing

In Chapter 4 findings from the descriptive data analysis were reported. This chapter presents results of statistical data analysis and hypotheses testing. The first part of this chapter provides a description of the research questions, main hypotheses and sub-hypotheses. The second part presents the results of tests arranged by main variables, namely, user characteristics, topic characteristics, search process characteristics and interface usability characteristics.

5.1 Main and sub-hypotheses

The prime objective of the present investigation was to evaluate how and to what extent a thesaurus-enhanced search interface can assist end-users in selecting search terms for query expansion. Specifically, it attempts to ascertain users' attitude toward both the thesaurus and interface as tools for facilitating search term selection for query expansion. It also intends to identify searching and browsing behaviours of users interacting with a thesaurus-enhanced interface attached to a large bibliographic database.

Based on this objective key main research questions were developed as follows:

- Are there any common patterns of user behaviour in thesaurus-based browsing and searching?
- What relationships are there between users' initial query terms and the terms they selected from the thesaurus for query expansion?
- Does topic complexity affect user-thesaurus interaction in general and search term selection in particular?
- Does topic familiarity affect user-thesaurus interaction?
- Does interface usability affect thesaurus browsing or other search behaviours?

Five main hypotheses relating to five categories of variables were formulated to address the research questions. Main hypotheses were then broken down into sub-hypotheses in order to specifically address and test variables entailed in each main hypothesis, as listed in section 3.3.3.

It should be noted that while the sub-hypotheses have been grouped together under main hypotheses, each sub-hypothesis, depending on the variables employed, may test a range of specific hypotheses. As mentioned in section 3.7.5 of the methodology chapter, the outliers have been removed to ensure the data normalisation. Four cognitive and two physical move outliers at the search level and three physical and two cognitive outliers at the user level have been identified and removed.

5.2 Hypothesis Testing

5.2.1 User characteristics

In this section the effects of user variables such as subject knowledge, gender, status, and computer or scientific database search experience on search and browse behaviour were tested.

Main Hypothesis 1: User characteristics have an effect on searching and browsing behaviour

H1-1: Users with more subject knowledge i.e. Academic staff make more cognitive and physical moves.

A two-sample *t*-test was carried out to compare the cognitive moves made by academic staff and postgraduate researchers. The results are shown in Table 5.1.

Status	N	Mean cognitive moves	StDev	CI
Academic staff Postgraduates	13 15	29.1 29.5	5.2 6.2	(-4.9, 4.1)

Table 5.1 *t*-test results for cognitive moves by status

As Table 5.1 shows that no significant difference was found between academic staff and the postgraduate researchers in terms of the mean number of cognitive moves made.

In addition to testing the total of cognitive moves versus status, individual cognitive move types were also *t*-tested against status to ensure that the pooling of total cognitive moves did not skew the significance of tests. Table 5.2 shows mean numbers of individual cognitive moves by type for both academic staff and postgraduates. Only the difference between the

means for term selection moves was found nearly significant (p=0.067). This cognitive move type is concerned with the number of times a user made term selections. The finding suggests that the academic staff on average made more selections than did the postgraduates.

Cognitive moves	Mean	
	Academic staff	Postgraduates
Term input	2.8	2.4
Browsing terms in mapping state	2.4	2.1
Browsing terms in thesaurus hierarchy state	0.64	0.52
Selection of terms	3.0	2.5
Combine search terms	2.1	2.1
Browsing retrieved titles	1.1	1.2
Query reformulation	1.1	1.1

Table 5. 2 Mean numbers of individual cognitive moves by type

The second part of H1-1 was concerned with physical moves and their relation to the users' status. The *t*-test reported in Table 5.3 indicates that there was no significant difference between users' status and the number of physical moves they made.

Status	N	Mean physical moves	StDev	CI
Academic staff	13	58.3	14.1	(-11.01, 11.63)
Postgraduates	14	58.0	14.4	

Table 5.3 *t*-test results for physical moves by status

To provide a comparative view of physical moves made by academic staff and postgraduates, Table 5.4 provides means of individual physical move types.

Physical moves	Mean		CI
	Academic staff	Postgraduates	
Perform search	3	2.5	(-0.118, 1.027)
Scroll up & down	7.7	6.4	(-1.01, 3.60)
Back & forward	0.22	0.61	(-0.845, 0.072)
Continue	3.8	3.3	(-0.359, 1.268)
Combine	1.9	2.1	(-0.610, 0.338)
Citation display &e-mail	2.9	3.3	(-0.887, 0.024)
Main search page	0.11	0.09	(-0.1071, 0.1526)
Previous & next page	0.7	0.7	(-0.358, 0.494)
Search history	0.18	0.11	(-0.0833, 0.2197)
Expand & contract	0.3	0.3	(-0.259, 0.350)

Table 5.4 Mean numbers of individual physical moves by type

As can be seen from Table 4 there were no differences between academic staff and postgraduates in terms of the number of physical moves they made during the search process. However, there were two move types whose p-values were low. The number of moves associated with citation display and e-mail was higher in the postgraduate group. The p-value for this move type was nearly significant (p= 0.063). This finding suggests that postgraduates tended to view the complete record more often than did academic staff. The second move was concerned with the use of *back* and *forward* buttons. Again, postgraduates made more moves of this type during their searches. The p-value for this test was 0.098. The general trend for this move suggests that postgraduates tended to use this feature more than academic staff.

$H_{1\text{-}2}$: Academic staff and postgraduate researchers differ in the number of terms they browse and select.

A two-sample *t*-test was conducted to ascertain whether or not user status affects browsing behaviour. Table 5.5 shows the test and its details. The test indicated that there is a nearly significant difference (p=0.057) between the number of terms browsed by academic staff and

postgraduate researchers. The results show that academic staff on average browse 11 more terms per search than postgraduate researchers.

Status	Number Of searches	Mean	StDev	CI
Academic staff Postgraduates	15 15	130.8 96.2	55.7 37.3	(-1.1, 70.3)

Table 5.5 *t*-test results for status and number of search terms browsed

This finding could be attributed to the fact that academic staff with wider domain knowledge tend to browse and navigate the thesaurus more than postgraduate researchers.

To establish whether the status had any effect on the number of terms selected a *t*-test was carried out and test results are shown in Table 5.6. Although it was found that on average academic staff selected 0.7 more terms than did the postgraduates, the p-value was not significant (p=0.5), reflecting the fact that there was no significant difference between the number of terms selected by academic staff and postgraduate researchers.

Status	Number Of searches	Mean	StDev	CI
Academic staff Postgraduates	15 15	17 15.2	7.51 6.08	(-3.46, 6.79)

Table 5.6 *t*-test results for status and selection of search terms

H_{1-3}: There is no significant difference in cognitive or physical moves made by male and female users.

A *t*-test was carried out to establish whether gender had an effect on cognitive moves. The p-value for cognitive moves was found to be nearly significant (p=0.054). This difference indicates that on average female users made fewer cognitive moves than did male users. However, the p-value for physical moves was large (p=0.145) reflecting the fact that there was no significant difference between the number of physical moves made by female and male users. Table 5.7 shows that female users also made fewer physical moves than did male users, although this difference was not statistically significant. This finding provides weak evidence that gender may affect the number of cognitive moves a user makes but does not appear to affect the number of physical moves.

Gender	N	Mean cognitive	N	Mean physical
Female	16	27.4	16	54.6
Male	12	31.8	11	63.4

Table 5.7 Mean numbers of cognitive and physical moves by gender

$H_{1\text{-}4}$: Level of computer experience has an effect on users' cognitive and physical moves.

A *t*-test was conducted to evaluate the effect of computer experience on the number of cognitive and physical moves users made during the search process. The p-values for both cognitive (p=0.7) and physical moves (p=0.9) were found to be large reflecting the fact that computer experience did not have any effect on the number of cognitive and physical moves. Table 5.8 shows the mean for both cognitive and physical moves with respect to level of computer experience.

Computer Experience	N	Mean Cognitive	N	Mean Physical
Experienced	21	29.5	19	58.0
Novice	7	28.5	8	58.8

Table 5.8 Computer experience and cognitive and physical moves

5.2.2 Topic Complexity Characteristics

The topic-related variables were analysed to investigate possible relationships between topic characteristics and the user's search moves. The results are summarised in the succeeding sections.

Main hypothesis 2: Topic complexity has effects on user-thesaurus interaction

H2-1: Increased topic complexity leads to an increase in the number of cognitive and physical moves.

To evaluate the effect of topic complexity on cognitive moves made during the search process a *t*-test was conducted. Since outliers have an undue influence on the general distribution of data, four searches identified as outliers in terms of their number of cognitive moves were excluded to provide a normally distributed data set. As shown in Table 5.9 complex topic searches involved around six more cognitive moves per search than those involving simple topics. The *t*-test showed this difference to be highly significant ($p < 0.001$), and the confidence interval (CI) indicates that there is a 95% probability that the difference lies in the range of 4.9 to 7.9 moves.

Complexity level	N	Mean Cognitive	CI
Simple topics	51	9.9	
Complex topics	35	16.2	(4.9, 7.9)

Table 5.9 Mean numbers of cognitive moves by topic types

The difference between the mean number of physical moves with respect to simple and complex topics was also compared using a *t*-test. Table 5.10 shows that complex topic searches were associated with an average of around seven more physical moves per search than for simple topics. Once again the difference was highly significant ($p<0.001$) and the 95% Confidence Interval indicates the typical range of values this difference would fall within.

Complexity level	N	Mean physical	CI
Simple topics	51	17.3	
Complex topics	37	24.7	(4.4, 10.5)

Table 5.10 Mean numbers of physical moves by topic types

$H_{2\text{-}2}$: Searches with complex topics require more query reformulation than those with simple topics

Table 11 shows the result of *t*-test for two variables namely topic complexity and query reformulation. The p-value was found to be highly significant ($p<0.001$). As can be seen from Table 5.11 complex topics were reformulated approximately twice as many times as were simple topics. This finding provides evidence in support of the above hypothesis that within a thesaurus-enhanced search environment complex queries

are subject to more query reformulation than searches involving simple topics.

Topic complexity	N	Mean query reformulation	CI
Complex topics Simple topics	39 51	1.5 0.88	(0.30, 0.96)

Table 5.11 Topic complexity and query reformulation

H2-3: Searches with complex topics are associated with the selection of more search terms.

A *t*-test was carried out to establish whether complex topics were associated with the selection of more search terms (Table 5.12). The p-value was found to be highly significant ($p<0.001$) indicating evidence to support the above hypothesis. On average, there were four more terms selected with respect to complex topics.

Topic complexity	N	Mean	CI
Complex topics Simple topics	39 51	7.7 3.5	(3.0, 5.4)

Table 5.12 Complex topics and term selection

5.2.3 Topic Familiarity, Prior Topic Search Experience and Search Type

This section provides the results of tests carried out on topic familiarity, topic search experience, search type and their relationship with cognitive and physical moves.

Main hypothesis 3: *Topic familiarity and search type have an effect on searching and browsing behaviour.*

$H_{3\text{-}1}$: Topic familiarity, topic search experience and search type affect cognitive and physical moves.

As in $H_{2\text{-}1}$ the number of search topics, excluding outliers, and mean number of cognitive and physical moves for each level of topic familiarity are shown in Table 5.13.

Level of topic familiarity	Cognitive moves		Physical moves	
	Mean	N	Mean	N
Unfamiliar	10.8	14	16.2	15
Moderately familiar	13.0	49	21.7	50
Very Familiar	12.3	23	20.5	23

Table 5.13 Moves by topic familiarity

A one-way ANOVA test indicated that there was no significant difference between topic familiarity levels in either the number of physical ($p=0.07$) or cognitive ($p= 0.3$) moves. However, the trend in the data indicates that topics identified as being moderately or very familiar were associated with around two more cognitive and five more physical moves than those topics identified as unfamiliar.

The effects on cognitive and physical moves of prior topic search experience and search type were examined. The results are shown in Tables 5.14 and 5.15.

Topic search experience	Cognitive moves		Physical moves	
	Mean	N	Mean	N
Experienced	12.7	55	20.8	55
No experience	12.1	31	19.08	33

Table 5.14 Moves by previous topic search experience

Search type	Cognitive moves		Physical moves	
	Mean	N	Mean	N
Broad	11.9	54	21.3	55
Specific	13.4	32	19.0	33

Table 5.15 Moves by search type

A *t*-test was carried out to confirm that there was no relationship between prior topic search experience and cognitive or physical moves. The p-values for cognitive and physical moves were 0.59 and 0.58 respectively indicating that topic search experience did not affect the number of cognitive and physical moves. The same test also indicated no significant difference between the means of cognitive or physical moves for broad and specific search types. The p-values for cognitive and physical moves with respect to broad and specific topics were 0.18 and 0.17 respectively. These findings suggest that neither search type nor previous topic search experience affected the number of moves users made during the search process. However, the general trend in the data shows that when conducting specific searches users made around two more cognitive and two less physical moves than did they when performing broad searches.

$H_{3\text{-}2}$: Topic familiarity affects the number of search term browsed and selected

A one-way ANOVA test was conducted to evaluate the relationship between topic familiarity and the number of browsed and selected terms. The results are shown in Table 5.16. The p-value for topic familiarity and browsed terms was significant (p=0.027) indicating the fact that the users who were more familiar with the topic they were searching for browsed more search terms than did those who had a low degree of familiarity. Table 5.16 shows that users in the moderately and very familiar categories browsed on average 17 and 19 more search terms respectively. This finding also demonstrates that those who have more

subject knowledge tend to explore a terminological space such as a thesaurus more than those with less subject knowledge.

Topic familiarity	Browse		Select	
	Mean	N	Mean	N
Unfamiliar	23.5	15	4.8	15
Moderately familiar	40.3	52	5.7	52
Very familiar	42	23	4.9	23

Table 5.16 Topic familiarity and browsed and selected terms

However, the ANOVA test for topic familiarity and the number of terms eventually selected indicated no significant difference between topic familiarity levels in the number of selected terms. An explanation for this could be the fact that those with a high level of familiarity tend to select only relevant and specific terms while browsing a larger number of terms.

H3-3: Topic search experience has an effect on the users' judgment of additional terms provided by the thesaurus.

The users' judgement of the additional terms was measured using three variables: whether the thesaurus provided additional terms, whether those terms were useful and how close were those terms to the original search topic. A chi-square test was carried out to evaluate the relationship between previous topic search experience and the provision of additional terms by the thesaurus. The results are shown in Table 5.17. The p-value was significant (p=0.005) reflecting the fact that topic search experience had an effect on users' judgement as to whether the thesaurus provided them with additional terms. This finding suggests that those who had not searched the topic previously indicated that they had found more additional terms than did those with previous topic search experience.

Topic search experience	No additional Terms	Additional terms	Total
Experienced	15	42	57
No experience	1	32	33

Table 5.17 Topic search experience and additional search terms

The relationship between topic search experience and usefulness of the additional terms was evaluated using a chi-square test. The p-value was found to be significant (p=0.012) indicating that topic search experience had an effect on users' evaluation of usefulness of additional terms they selected from the thesaurus. Table 5.18 shows the results of the test. Those with no topic search experience judged the additional thesaurus terms to be more useful than did those with prior topic search experience.

Topic search experience	Term usefulness		Total
	Term not useful	Term useful	
Experienced	4	29	33
No experience	21	36	57

Table 5.18 Topic search experience and term usefulness

A chi-square test was carried out to evaluate whether topic search experience had an effect on users' judgement on the closeness of additional terms to the original topic. The p-value was found to be significant (p=0.021). This significance indicates that users with no previous topic search experience found additional terms closer to their original search topics than did those with previous topic search experience.

5.2.4 Search Process Characteristics

Results of statistical tests associated with search process variables namely cognitive and physical moves and their effect on result relevance, term usefulness and search result satisfaction are reported in this section.

Main hypothesis 4: Search process characteristics affect search term and search result characteristics

$H_{4\text{-}1}$: There is a relationship between result relevance and cognitive and physical moves.

A two-sample *t*-test was conducted to establish whether there was any relationship between cognitive moves and the relevance rating of result sets by users. The p-value (p=0.22) indicated there was not sufficient evidence to support this hypothesis. Indeed as Table 5.19 shows searches in the category of *very relevant* are associated with slightly

Relevance	Cognitive		Physical	
	N	Mean	N	Mean
Partially relevant	27	13.3	29	23.4
Very relevant	59	12	59	19

Table 5.19 Cognitive moves and result relevance

fewer cognitive moves than those in the *partially relevant* category. Thus, simply making more cognitive moves dose not guarantee the retrieval of very relevant document sets.

The same test was carried out for the relationship between physical moves and the relevance ratings of result sets (Table 5.19). The p-value was found to be significant (p=0.015) indicating that searches whose result sets were evaluated as *very relevant* were associated with fewer physical moves than those which were evaluated as *partially relevant*. This finding may have two implications. First, it may indicate that poor

results led the users to more wandering around and making more physical moves. Second, it may reflect the significance of the quality of physical moves rather than their quantity.

$H_{4\text{-}2}$: Cognitive and physical moves affect users' satisfaction with the number of records retrieved.

One-way ANOVA tests were carried out for both cognitive and physical moves with respect to users' satisfaction with the number of records retrieved. Neither of the tests showed any significance indicating that user satisfaction with the number of retrieved records is not dependent on the number of search moves. Therefore, this hypothesis is rejected.

$H_{4\text{-}3}$: There is a relationship between cognitive and physical moves and users' evaluation of term usefulness.

A two-sample *t*-test of cognitive moves and term usefulness was conducted and the results are shown in Table 5.20. There was a significant difference (p=0.043) between cognitive moves made by those who had evaluated the selected terms as useful and those who did not. On average just over two more cognitive moves were associated with those searches whose selected terms were evaluated as being useful.

Term usefulness	**Cognitive**		**Physical**	
	N	Mean	N	Mean
Useful	62	13.1	63	21.0
Not useful	24	10.8	25	19.0

Table 5.20 Cognitive and physical moves and term usefulness

The same test was conducted for physical moves and term usefulness. However, the p-value (p=0.28) indicated that physical moves did not have any relationships with users' evaluation of term usefulness. These

findings point to the fact that while cognitive moves are important in satisfying users with their search term selection experience using a thesaurus, physical moves are not.

$H_{4\text{-}4}$: There is a relationship between the number of selected terms and users' evaluation of term usefulness.

A two-sample *t*-test showed a significant relationship (p=0.006) between the number of selected terms and term usefulness. As Table 5.21 shows those who evaluated their search terms as being useful selected on average around two more search terms than did those who did not evaluate the terms as being useful.

Status	N	Mean	CI
Useful Not useful	65 25	5.8 4	(-3.035, -0.528)

Table 5.21 Number of selected terms and term usefulness

$H_{4\text{-}5}$: There is a relationship between the number of selected terms and users' judgement of the relevance of result sets.

The relationship between the number of selected terms and the relevance evaluation of retrieved sets was assessed using a *t*-test and the results are shown in Table 5.22.

Status	N	Mean	CI
Very relevant Partially relevant	60 30	4.8 6.4	(-0.062, 3.162)

Table 5.22 Number of selected terms and result relevance

The p-value was found to be marginally significant (p=0.06) indicating that there may exist a relationship between the number of search terms users

select and the way they evaluate the relevance of the retrieved set. As can be seen from Table 5.22 searches whose results were judged as being very relevant were associated with fewer selected terms whereas those searches whose results were evaluated as being partially relevant were associated with more selected terms. Searches in the *partially relevant* category had an average of two more search terms selected than those in the *very relevant* category. This finding indicates that the quantity of search terms is not the prime determinant of the relevance evaluation of retrieved records.

$H_{4\text{-}6}$: Users who spend more time looking at search terms make more cognitive moves.

It should be noted that unlike previous hypotheses this particular hypothesis was tested at the user level rather than search level. The reason for this lies in the fact that time has been considered as a search process variable in this study, hence this hypothesis has been coupled with other search process related hypotheses.

The time spent for entering a search term, browsing the mapped terms and selection of terms and the users' cognitive moves. A *t*-test was run to compare the mean number of cognitive moves of users who spent less than or more than 10 minutes in these search activities. In order to undertake this analysis, two categories of users were identified; those who spent 10 or less than 10 minutes on search terms and those who spent more than 10 minutes. Table 5.23 shows cognitive moves and time spent on search terms. The test shows a significant difference ($p=0.029$) indicating that users who spent more time on browsing and selecting search terms tend to make more cognitive moves.

Time	N users	Mean cognitive	StDev	CI
Spending less than 10 minutes Spending more than 10 minutes	19 9	27.4 33.1	4.67 5.99	(-8.78, -0.67)

Table 5.23 *t*-test for Cognitive moves and searches time spent for search term

Physical moves were not evaluated against time, as it is self evident that more physical moves make users to spend more time.

5.2.5 Usability Characteristics

This section provides the results of statistical tests on the usability of the search interface and the thesaurus.

Main hypothesis 5: Perceptions of interface usability have an effect on searching and browsing behaviour.

H5-1: There is a significant relationship between the ease of learning and using the interface and ease of thesaurus browsing.

A chi-square test was conducted to evaluate the relationship between the ease of learning and use of the interface and ease of thesaurus browsing. The p-value was found to be significant (p=0.007) indicating the fact that those who evaluated the learning and use of the interface as *very easy* also judged the thesaurus browsing and navigation to be *very easy*. Table 5.24 shows results of the *t*-test.

Ease of learning and use of interface	Ease of thesaurus browsing/navigation		Total
	Easy	Very easy	
Easy Very easy	16 4	3 7	19 11
Total	20	10	30

Table 5.24 Ease of learning and use of interface and ease of thesaurus browsing/navigation

This finding indicates that search interface usability is associated with thesaurus browsing/navigation and the more usable a search interface is, the easier thesaurus browsing will be.

$H_{5\text{-}2}$: There is a significant relationship between ease of thesaurus browsing/navigation and cognitive and physical moves.

All users' judgements on thesaurus interface usability fell into two categories namely *easy* or *very easy* on a Likert scale. Therefore, a two-sample *t*-test was conducted to test the relationship between ease of thesaurus browsing and physical moves users made during their search process. The p=value for this test was found to be highly significant (p=0.003) reflecting the fact that users who evaluated thesaurus browsing as *very easy* made fewer physical moves than did those who had evaluated the ease of browsing as *easy*.

Ease of thesaurus browsing	**N**	**Mean moves**	**CI**
Easy	16	64.1	(5.7, 23.7)
Very easy	11	49.4	

Table 5.25 Ease of thesaurus browsing and physical moves

As table 5.25 shows users in the *easy* category made on average 15 more physical moves. This finding suggests that those who have evaluated their thesaurus browsing experience as easy had to make more use of system features to perform their tasks compared to those who found the browsing *very easy*.

The same test was carried out to evaluate the relationship between ease of thesaurus browsing and cognitive moves. The p-value (p=0.25) indicated that there was no relationship between these two variables.

However, as the general trend in the data (Table 5.26) shows on average users who judged the thesaurus browsing experience *easy* made on

Ease of thesaurus browsing	N	Mean moves	CI
Easy Very easy	17 11	30.2 27.8	(-1.8, 6.7)

Table 5.26 Ease of thesaurus browsing and cognitive moves

average around two more cognitive moves than did those who evaluated it as *very easy*. These findings suggest that while physical moves have direct relationship with the way in which users evaluated thesaurus browsing/navigation, cognitive moves did not appear to have any effect on users' judgement of thesaurus browsing.

H5-3: There is a significant relationship between ease of thesaurus browsing/navigation and the number of terms browsed and selected.

A *t*-test was conducted for this hypothesis and the results are shown in Table 5.27.

Ease of thesaurus browsing	N	Mean terms browsed	Mean terms selected
Easy Very easy	19 11	125.1 93.5	16.5 15.5

Table 5.27 Ease of thesaurus browsing and number of terms browsed and selected

The p-value for testing the relationship between the ease of thesaurus browsing and the number of terms browsed was found to be nearly significant (p=0.06) indicating the fact that users who evaluated the ease of thesaurus browsing/navigation as *easy* browsed a larger number of

search terms than did those who judged their browsing experience as *very easy*. Ease of thesaurus browsing/navigation did not have any effect on the number of terms users browsed.

The p-value for ease of thesaurus browsing and the number of selected terms was large (p=0.68) indicating no significant difference between the mean number of terms selected by users in both *easy* and *very easy* categories. The general trend in Table 5.27 suggests that users who judged thesaurus browsing a *very easy* task tended to browse and select fewer search terms compared to those who felt that browsing the thesaurus was less easy.

$H_{5\text{-}4}$: There is a significant relationship between ease of thesaurus browsing/navigation and the number of query reformulation instances.

A t-test was conduct to find out if there is any statistically significant relationship between ease of thesaurus browsing and the number of query reformulations carried out by users. The p-value was found to be significant (p=0.013) reflecting the fact that there exists a meaningful relationship between the two variables. Table 5.28 shows the results of the test.

Ease of thesaurus browsing	**N**	**Mean query reformulation**	**CI**
Easy Very easy	19 11	1.6 0. 55	(0.24, 1.83)

Table 5.28 Ease of thesaurus browsing/navigation and query reformulation

Users who evaluated thesaurus browsing as *easy* tended to reformulate their queries more than those who judged thesaurus browsing as *very easy*. This finding reflects that users with more reformulation instances found thesaurus browsing less easy than did those with fewer

reformulations. The explanation for this could be that query reformulation challenged users and had more cognitive load on the part of users to interact with the thesaurus for reformulation and as a result they evaluated thesaurus browsing less easy.

H5-5: There is a significant relationship between the ease of thesaurus browsing/navigation and the time spent per search.

A t-test was conducted to evaluate whether or not there was any relationship between the time users spend on their searches and the ways in which they evaluated thesaurus browsing. The p-value was found to be statistically significant (p=0.016) indicating that users who evaluated thesaurus browsing *very easy* spent less time on their searches than did those who evaluated it as *easy*.

Ease of thesaurus browsing	N	Mean Time (min)	CI
Easy Very easy	19 11	23.4 16.8	(1.3, 11.8)

Table 5.29 Ease of thesaurus browsing and time spent on searches

As Table 5.29 shows that users in the category of *very easy* spent on average over six minutes less on their searches than did those who evaluated thesaurus browsing as *easy*. This finding suggests that the easier the users found thesaurus browsing, the faster they tended to complete their searches.

5.3 Summary

Table 5.30 summarises the results of statistical tests reported earlier together with an indication as to whether a given hypothesis is proven or rejected, whether it corroborates or contradicts the findings reported in the literature or whether it is a novel finding.

P=Proven R=Rejected Cor= Corroborated Con=Contradicted N=Novel

H	Hypothesis	P	R	Cor	Con	N
1-1	Users with more subject knowledge (i.e. Academic staff) make more cognitive and physical moves		✓			✓
1-2	Academic staff and postgraduate researchers differ in the number of terms they browse and select	✓				✓
1-3	There is no difference in cognitive or physical moves made by male and female users		✓			✓
1-4	Level of computer experience has an effect on users' cognitive and physical moves		✓	✓		
2-1	Increased topic complexity leads to an increase in the number of cognitive and physical moves	✓		✓		
2-2	Searches with complex topics require more query reformulation than those with simple topics	✓		✓		
2-3	Searches with complex topics are associated with the selection of more search terms.	✓				✓
3-1	Topic familiarity, topic search experience and search type affect cognitive and physical moves.		✓		✓	
3-2	Topic familiarity affects the number of search term browsed and selected	✓				✓
3-3	Topic search experience has an effect on the users' judgment of additional terms provided by the thesaurus.	✓				✓
4-1	There is a relationship between result relevance and cognitive and physical moves.		✓	✓		
4-2	Cognitive and physical moves affect users' satisfaction with the number of records retrieved.		✓	✓		
4-3	There is a relationship between cognitive and physical moves and users' evaluation of term usefulness.	✓				✓
4-4	There is a relationship between the number of selected terms and users' evaluation of term usefulness.	✓				✓
4-5	There is a relationship between the number of selected terms and users' judgement of the relevance of result sets.	✓				✓
4-6	Users who spend more time looking at search terms make more cognitive moves.	✓				✓
5-1	There is a relationship between the ease of learning and using the interface and ease of thesaurus browsing.	✓				✓
5-2	There is a relationship between ease of thesaurus browsing/navigation and cognitive and physical moves.	✓				✓
5-3	There is a relationship between ease of thesaurus browsing/navigation and the number of terms browsed and selected.	✓				✓
5-4	There is a relationship between ease of thesaurus browsing/navigation and the number of query reformulation instances.	✓				✓
5-5	There is a relationship between the ease of thesaurus browsing/navigation and the time spent per search.	✓				✓

Table 5.30 Summary of statistical test results

This chapter presented the results obtained from the statistical tests of hypotheses developed within this research. Five categories of

hypotheses were tested: user related hypotheses, topic complexity hypotheses, topic familiarity hypotheses, search process hypotheses and the interface and thesaurus usability hypotheses. The next chapter will suggest interpretations and explanations for the findings reported in this chapter together with those in Chapter 4.

Chapter 6: Discussion

The purpose of this chapter is to discuss and provide interpretations relating to the findings from the descriptive data analysis in chapter 4 and the statistical analysis and hypothesis testing in chapter 5. The format of this chapter is based on the main hypotheses whose test results were reported in the previous chapter. Descriptive findings relating to each main hypothesis have also been discussed where appropriate. The findings have been discussed with reference to existing literature.

6.1 User characteristics

The present investigation involved two user groups namely academic staff and postgraduate students. The effect of a variety of user characteristics on the search process was evaluated. It was revealed that some user characteristics are likely to affect users' cognitive and physical move patterns.

The findings suggest that the groups differ in terms of their use of certain cognitive move types. Academic staff browsed on average 11 more thesaurus terms per search than did postgraduates. Another of the

cognitive move types was the number of instances a user made a term selection with academic staff members making more selections than did postgraduates.

The finding that academic staff members browsed more search terms and made more selections suggests that users with more domain knowledge tend to explore thesaurus terms more than users with a low level of subject knowledge. The explanation for this could be that academic staff members are more familiar with their subject area and browsing terms provide them with an opportunity to more specifically formulate their queries. This finding concurs with Vakkari (2002) who stated that people with more advanced subject knowledge can identify more variant expressions of a concept and select more pertinent terms for expressing facets among matching thesaurus terms.

However, users' status did not have any significant effect on the number of physical moves i.e. their use of system features. This was due to the necessity of using certain system features to progress through a search. For instance, features such as *continue* or *scroll up and down* during thesaurus browsing or viewing results were typical and all users had to make use of them. Of the physical moves only two moves namely *citation display* and *back and forward* were found to be marginally affected by status. Postgraduates on average made more *citation display* moves, which are concerned with viewing full records, than did academic staff members. This finding could be explained by the fact that postgraduates with limited domain knowledge felt the need to view the full bibliographic record in order to decide on the usefulness and relevance of records. The *back and forward* moves are associated with the use of features in Internet browsers. Postgraduates made this move three times as often as academic staff members. The interpretation for this finding is that postgraduates had to make more moves around different search states such as browsing, selection of terms and viewing full records. The reason

for this might be due to a lack of certainty and their struggle for double-checking the selections they have made at previous stages. This finding can be compared with those reported by Bilal and Kirby (2002) who found that children made more backtracking moves (i.e. the use of Back button in the Netscape browser) than did graduate students in searching a Web search engine.

Since the Boolean operators were explicitly displayed at both term selection and search combination stages users made use of them in most of their searches. All thirty users introduced the AND operator while 21 users made use of the OR operator.

The effect of computer experience was studied as one of the independent variables in user characteristics. The effect of computer experience on the number of cognitive and physical moves users made was assessed. It was found that neither cognitive nor physical moves were dependent on user's computer experience. This finding can be compared with those reported by (Borgman et al., 1995) who found that computer experience had minimal effect on children's searching behaviour while conducting searches in an online library catalogue.

The relationship between the time users spent on search term selection and the number of cognitive moves they made was assessed. It was found that users who spent more than ten minutes formulating their search terms made on average five more cognitive moves than did those who spent less than ten minutes. This finding indicates that in a thesaurus-aided search environment, a considerable amount of users' time is spent on browsing and/or selecting terms. This finding also suggests that some aspects of user interface interaction can be predicted based on the time users spend in their initial term selection state. This may help identify behavioural patterns subsequent to the search term selection stage.

It should be noted that cognitive and physical moves were treated and counted as being equivalent to the user and no value judgment or weighting of different types of move was conducted. However, within the context of a thesaurus-enhanced search interface some move types can be viewed as more important than others. For instance thesaurus browsing or term selection may be viewed as more important than combination of search terms or viewing search results.

Gender has been found to have no (Qiu, 1993) or minimal effect (Borgman et al., 1995) on the search process or search outcomes. The role of individual differences such as gender is important in studying different patterns of search behaviour, in particular, search term selection. The effect of gender differences on search behaviour can be evaluated and fed into the improvement of users' and systems' performance. In this study, it was found that gender had a nearly significant effect on the number of users' cognitive moves, while no significant effect of gender was found on physical moves. The explanation for this may lie in the fact that the female users selected fewer search terms than did male users. Therefore, the smaller number of terms selected by female users led them to make fewer cognitive moves. Test of gender with the variables search time and number of records browsed showed significant differences. Female users on average spent less time on their searches and browsed fewer records than did male users. However, search outcome measures namely user satisfaction with the number of results and the number of terms indicates a different trend. It was found that female users were significantly more satisfied with the number of terms they selected and the number of results they retrieved. An interesting observation was the fact that female users on average made fewer physical moves than did male users, though the difference was not statistically significant. Considering the number of physical moves and user satisfaction with terms and results it can be suggested that female users were more efficient in their searches than were male users. These

findings can be contrasted with those reported by Ford and Miller (1996) and Ford et al. (2001) who studied the role of individual differences such as gender and age in Internet searching. They reported that female gender was associated with poor retrieval performance and that females reported: being unable to find their way effectively around the Internet, getting lost, not being in control, and only looking at things suggested to them.

6.2 Topic Complexity

Research has shown that the complexity of topic searches affects information retrieval (Kristensen and Jarvelin, 1998; Saracevic et al., 1988) and that more complex topics require a higher level of engagement and effort in an interactive searching environment (Fowkes and Beaulieu, 2000).

Unlike the above studies, the focus of the present investigation was to study the effect of topic complexity on the search process and users' interaction with the system rather than the search outcome. In this study complex topics were found to involve significantly more cognitive and physical moves compared to simple topics. Although the notion of complexity was defined slightly differently, the findings are consistent with those reported by Marchionini (1989) who found that the number of moves was dependent on the task and that more complex tasks lead to more moves being made during the search and cause users to make more use of system features.

Another interesting finding concerned the effect of topic complexity on users' evaluation of the usefulness of thesaurus suggested terms. It was found that users who searched for complex topics evaluated the thesaurus suggested terms as being more useful than did those with simple topics. This finding may indicate that complex topics benefit from thesaurus terms more than simple topics.

The effect of topic complexity on query reformulation was also examined. It was found that query reformulation instances were significantly associated with complex topics. This finding concurs with those reported by Fowkes and Beaulieu (2000) who found that query reformulation and interactive query expansion appeared to be beneficial for complex topics which required a higher level of engagement and effort.

An analysis of the number of query expansion terms also showed that complex topic queries were expanded with twice as many terms as were simple topic queries. This finding suggests that in a thesaurus-enhanced search environment complex topics are subject to more query expansion and modification than simple topics.

In addition to the effect of topic complexity on cognitive and physical moves, a significant difference was found between simple and complex topic searches in the number of terms selected. Complex topic searches had on average four more terms selected than did simple topics. This finding suggests that complex topics are associated with more selected search terms. One possible explanation for this finding could be that the thesaurus provided complex topic searches with more chances of term selection through presenting alternative search terms than it did for simple topics.

6.3 Topic Familiarity

Research has shown that topic familiarity affects some information search behaviours. Kelly and Cool (2002) suggest that the more familiar a person is with a search topic, the less time they spend reading retrieved documents. Hsieh-Yee (1993) found that subject knowledge affected the searching behaviour of experienced searchers. Experienced searchers who were not familiar with the topic relied more on the thesaurus for term

suggestion, made more effort in preparing for the search and included more synonyms and tried more combinations of terms, whereas experienced searchers familiar with the topic used more of their own terms.

The results from this study indicate that topic familiarity has an effect on users' behaviour while browsing terms in a thesaurus enhanced search environment. It was shown that moderately familiar and very familiar topics were associated with browsing around twice as many thesaurus terms as was the case for unfamiliar topics. This finding suggests that the more familiar a person with a topic, the more terms they tend to browse. This finding indicates that a high level of topic familiarity encourages end-users to explore a terminological space such as a thesaurus to specifically formulate their queries. This is in agreement with Vakkari (2002) who suggested that users with more subject knowledge have a larger and more specific vocabulary and tend to identify more variant expressions of a concept. The finding that users with a higher level of familiarity browsed more thesaurus terms and tended to make more physical and cognitive moves can be contrasted with that reported by Hsieh-Yee (1993) who concluded that experienced searchers who are familiar with a topic rely simply on their own terms for the search and make less effort finding additional terms. The explanation for this difference may lie in the fact that the subjects in Hsieh-Yee's study were professional searchers who were more equipped with search terms while those recruited for this study were end-users. An alternative explanation could be due to the difference of user interaction with the thesaurus for term selection. The users in Hsieh-Yee's study only had the opportunity to use a printed thesaurus whereas in this study the use of a thesaurus was part of the search process and all users interacted with the thesaurus.

Topic familiarity did not have any significant effect on the number of terms ultimately selected. A possible explanation for this finding could be that users with a high degree of topic familiarity tend to select only a few relevant and specific terms despite the fact that they browse a larger number of terms. These findings indicate that some search behaviours such as thesaurus browsing and term selection could be used to understand the extent to which users are familiar with a topic for which they wish to search. The finding could also contribute to the classification of users based on their familiarity level and to the suggestion of ways in which personalised interaction with a thesaurus-enhanced search system could be designed.

The effect of topic familiarity on cognitive and physical moves was not found to be significant. However, the results from this study indicate that topics identified as being *moderately familiar* or *very familiar* were associated with around two more cognitive and five more physical moves than those topics identified as being *unfamiliar*. This trend suggests that users with a high degree of familiarity tend to make more moves than do those with low familiarity levels. The finding that cognitive and physical moves may show a discrepancy in relation to topic familiarity is in line with the findings reported by Kelly and Cool (2002) who stated that some information search behaviours such as reading time and efficacy, which they define as the ratio of the number of saved documents to the total number of viewed documents, vary with respect to topic familiarity.

Another interesting finding was the fact that there was no relationship between topic familiarity and type of search. This can be compared with Vakkari (2002) who stated that the more familiar a user is with the domain, the more specific the search will be. An additional finding was that topic familiarity was unrelated to topic complexity. This finding suggests that whether or not the user is familiar with a topic does not

relate to the level of complexity of his search topic. This finding can also be contrasted to the case of expert searchers reported in Hsieh-Yee [6] who formulated more complex queries when conducting searches on topics with which they were not familiar. In her study, it was found that topic familiarity was related to query complexity i.e. searchers with low level of topic familiarity tended to formulate more complex queries. This discrepancy may be explained by the fact that in this study end-users were involved while in Hsieh-Yee's both expert searchers and end-users were recruited and that expert searchers may represent a different pattern of search behaviour.

6.4 Topic Search Experience and Search Type

Findings on the relationship between prior topic search experience, search type (broad/specific) and cognitive and physical moves suggest that neither prior topic search experience nor the search type had any significant effect on the number of cognitive or physical moves. However, it is interesting to note that users with specific topics tended to make two more cognitive and two less physical moves than did those who indicated a broad search. It was also found that users who stated that they were carrying out a specific search introduced significantly more initial terms than did those with broad search requests. This finding suggests that users who choose to perform a specific search hold a richer set of terms to introduce for query formulation than those selecting broad searches. It also indicates that users in this study could identify the difference between broad and specific searches and the fact that the more specific a search, the more initial search terms used.

Although prior topic search experience did not appear to affect the number of cognitive and physical moves, it did affect users' judgment of terms they selected from the thesaurus. Users who had no prior topic search experience found significantly more additional terms from the

thesaurus than did those who had prior topic search experience. Another interesting finding was the fact that users with no prior topic search experience evaluated thesaurus terms more useful than did those with previous topic search experience and found those terms closer to their original search topics. This finding reflects the fact that users with more knowledge of the topic tend to be more specific and distinctive. These findings are consistent with those reported by Vakkari (2000) who suggested that people with scarce domain knowledge need support for expanding and differentiating their conceptual model of the topic.

6.5 Search Term Characteristics

One of the main issues in relation to the selection of search terms is how the number of selected terms affects users' evaluation of result relevance and term usefulness. In this study it was found that those searches whose results were evaluated as being *very relevant* were associated with fewer search terms compared to those in the *partially relevant* category. This finding reflects the fact that quality of search terms rather than quantity plays a role in users' judgement.

Looked at from a different perspective the finding concerning term selection and term usefulness suggests that users who evaluated the thesaurus terms as being useful selected on average around two more search terms than did those who did not find the terms useful. This indicates that term satisfaction and the number of terms selected have a significant relationship with each other. An interesting observation in relation to users' evaluation of term usefulness was the fact that those searches in the category of *useful* were associated with significantly more cognitive moves. This finding suggests that users who make fewer cognitive moves evaluate their search terms to be more useful than those who make more cognitive moves.

Most studies on information retrieval behaviour evaluate the effect of search process variables such as search tactics and use of system features on search outcomes to assess the extent to which each contributes to successful searches (Borgman et al., 1996). In this study, the effect of cognitive and physical moves on the number of records retrieved and result relevance was assessed. An interesting finding was that there was a clearly significant relationship between physical moves and users' evaluation of the result set relevance. Searches whose results were evaluated as being *very relevant* were associated with around two less physical moves. The number of cognitive moves was not found to have any significant association with the relevance evaluation. However, the general trend in the data shows that those searches whose results were evaluated as being *very relevant* had two less cognitive moves than those evaluated as *partially relevant*. This finding reinforces the proposition that the quality of physical and cognitive moves is more important than their quantity. The finding that fewer physical and cognitive moves lead to *very relevant* result sets concurs with those reported by Marchionini (1989) who found that success is more likely to occur with a small number of moves.

6.6 Interface Usability

Research has been carried out into the usability of different types of thesaurus interfaces (Beaulieu, 1997; Sutcliffe et al., 2000b; Blocks et al., 2002). In this study, the effect of the search and thesaurus interface usability on search and browse behaviour was assessed. The finding on the relationship between the ease of learning/use of the interface showed that users who evaluated the process of learning and using the search interface as *very easy* also judged their experience of thesaurus browsing/navigation as *very easy*. This finding was substantiated by another observation concerning the time users spent on their searches. It

was shown that users who evaluated thesaurus browsing/navigation as being *very easy* spent less time on the searches than did those who evaluated it as simply *easy*.

The ease of thesaurus browsing/navigation was also found to have a significant relationship with physical moves. Users who found thesaurus browsing *very easy* made less physical moves. Users who evaluated thesaurus browsing/navigation as being *very easy* made less cognitive moves compared to those who judged it as being *easy*, though this difference was not found to be statistically significant. The general trend indicates that the fewer the cognitive and physical moves, the easier thesaurus browsing was perceived to be by users.

An interesting finding concerning the relationship between the ease of thesaurus browsing/navigation and the number of terms browsed was the fact that users who evaluated the thesaurus browsing/navigation as *very easy* browsed fewer thesaurus terms than did those who evaluated it as *easy*. A possible explanation for this could be that users who browsed a smaller number of terms faced fewer challenging situations and as a result found the thesaurus browsing experience easier than those who browsed a larger set of terms.

Despite the existence of a relationship between the number of terms browsed and the ease of thesaurus browsing, there was no significant relationship between the number of selected terms and the ease of thesaurus browsing. These findings suggest that the number of terms browsed varies with respect to the ease of thesaurus browsing.

Another finding concerning thesaurus browsing was the fact that users who made many query reformulations were found to evaluate the

thesaurus browsing/navigation to be less easy when compared to those who made fewer query reformulations. A possible explanation could be that query reformulation made the whole search process more challenging and users had to deal with a more cognitively intensive task in terms of thesaurus browsing/navigation and of making selections different from their original one. This may have resulted in thesaurus browsing/navigation being rated as *easy* rather than *very easy* by users with more query reformulations.

An interesting finding in relation to query reformulation was that around 76% of searches were expanded through adding thesaurus terms. This finding may reflect the fact that the subjects of this investigation, who were all end-users were able to make use of the thesaurus for selecting additional search terms. This finding also suggests that thesauri can be regarded as useful tools for end-users of a particular subject domain.

A reflection on the time spent per search and per search state and the number of screenshots viewed by users provide a picture of the whole search process from a different perspective. Users on average spent around half of their search time entering search terms, browsing thesaurus terms and selecting from those terms. They spent the other half combining search terms and viewing results. The screenshots viewed by users were roughly in proportion to time spent. On average users viewed nine screenshots relating to term input, browsing thesaurus terms and term selection. They viewed around seven screenshots per search while combining search terms and viewing results. This finding suggests that in a thesaurus-enhanced search setting where users are presented with opportunities to browse and select search terms, such search process characteristics as screenshots viewed and time spent per each search state can be viewed as dependent variables.

Another interesting finding with respect to search term screenshots was the fact that those who chose to perform a broad search viewed on average two less screenshots of search terms than did those who chose to carry out a specific search. An interpretation for this finding lies in the fact that on average broad searches had fewer initial terms than did specific searches.

6.7 Query Reformulation

Several studies have reported on a number of search term sources from which users may select search terms for query reformulation (Fidel, 1991a; Efthimiadis, 1996; Spink and Saracevic, 1997). In this study the sources of search terms during the query modification process were analysed. It was found that users relied on the thesaurus as the first source of search terms. Users' own terms were the second source and terms selected from the results retrieved were the third source. This finding can be compared with those by Spink and Saracevic (1997) who reported that the thesaurus was found to be the third source of search terms after the user's written request and user interaction. An explanation for this apparent difference may lie in the fact that the present study utilised a thesaurus-enhanced search setting with end-users interacting directly with the system whereas in their study thesauri were mainly used by professional searchers and the study took place in a mediated search environment. The current study suggests that thesauri, users' own terms and the retrieved results are the three most important sources of search term selection in an end-user searching environment. The users' preference in selecting terms from the thesaurus suggests that thesauri have the potential to assist end-users in the search term selection process.

Previous research has shown that semantic relationships in standard thesauri can have an impact on retrieval performance and can be used as

a source for query expansion (Greenberg, 2001a, 2001b). In the present study, users' query expansion behaviour within the context of a thesaurus-enhanced search environment was investigated. It was found that 76% of the searches were expanded and all users expanded at least one of their three queries.

Another finding concerned the search state at which the query expansion took place. Over 60% of searches were expanded at the thesaurus mapping state with the remaining 40% of searches being expanded at the thesaurus hierarchical state. An explanation for this finding lies in the fact that the thesaurus mapping state provides the initial list of thesaurus terms and most of the users made their decisions on selection of search terms at that stage. Only in around 30% of the searches did users investigate the context of individual thesaurus terms by moving on to the thesaurus hierarchical state and as a result end up selecting their search terms at that state. In fact, the availability of search terms at the earlier stage of the search encouraged users to select more search terms at the initial stage. This finding concurs with Beaulieu et al. (1995) who reported that the availability, number and quality of terms are factors affecting the query expansion process.

The finding that users selected their search terms at both thesaurus mapping and hierarchical states seems to indicate that in addition to the number and quality of terms presented, personal preference for investigating details of a thesaurus term was also an issue. This finding suggests that the search interface design should cater for the preferences of different types of users.

User-based interactive query expansion research has shown that the majority of query expansion terms chosen by users can be represented as hierarchical and associative relationships to the initial terms (Efthimiadis, 2000). Greenberg (2001) found that synonymous and

broader terms were perceived by end-users to have the greatest value as additional search terms. This study investigated the relationships between initial terms and query expansion terms selected by users. Around 30% of the expansion terms selected from the thesaurus were narrower terms. Related terms accounted for 28% of terms and synonymous terms a further 10%. Broader terms constituted 10% of expansion terms. The remaining 22% accounted for those expansion terms which were identified as having statistical relationship to users' initial terms. The finding that 78% of query expansion terms fall within hierarchical or associative relationships can be compared with Efthimiadis (2000) who reported that approximately 66% of the term associations were represented by hierarchical or associative relationships.

There are two findings that reflect the importance of the role of statistical techniques in suggesting thesaurus descriptors to the user. The first finding indicated that around 18% of search terms entered by users were statistical matches i.e. matched those descriptors which had frequently co-occurred with terms provided by users in the database records. The second finding showed that over 22% of query expansion terms selected by users were terms that were suggested to the user as a result of a statistical analysis. These two findings point to the fact that the use of thesauri can be enhanced through statistical analysis techniques.

Users in this study showed a strong tendency towards hierarchical relationships by selecting 40% of their expansion terms from either narrower or broader terms. It is also interesting to note that users chose 28% of their expansion terms from related terms. These findings could be contrasted to Jones et al. (1995) who found that the majority of terms retrieved by thesaurus navigation came through associative relationships. This apparent difference may be explained by the coverage and treatment of hierarchical and associative relationships in the thesauri used. The INSPEC thesaurus, which was used by Jones et al., contained

10,200 related term links and 5,995 hierarchical relationships. The CAB thesaurus, which was utilised in this study, provides a rich set of narrower and related terms. The 1999 edition of CAB thesaurus contains 15,617 RTs, 44,398 NTs and 12,156 BTs. The thesaurus is strong in suggesting narrower terms relating to organisms, taxonomy of species and animal diseases. Therefore, users were supported by a large vocabulary of terms for query expansion. The finding that a considerable number of query expansion terms were selected from narrower terms may simply reflect the fact that this type of relationship in the CAB thesaurus is particularly well-developed.

As reported earlier 22% of query expansion terms were identified as having a statistical relationship to users' initial terms. An example of a statistical relationship is as follows: when a user enters *septic arthritis* a list of suggested thesaurus terms including *septic arthritis* and *horses* appears. While both *horses* and *septic arthritis* are thesaurus descriptors, there is no thesaurus type relationship between these two terms. The Ovid search system uses a statistical algorithm for establishing relationships between the descriptors based on their co-occurrence in the documents. The finding that one fifth of query expansion terms were those resulted from statistical analysis of thesaurus descriptors indicates that there are alternative ways in which thesauri can be used to enhance thesaurus-based term suggestion facilities.

To look into expanded searches from a different perspective, it is helpful to consider users' impression on the provision of additional terms by thesaurus and the usefulness of those terms. Users in 82% of searches commented that the thesaurus provided additional terms. It was also stated that in 87% of these cases the additional terms provided by the thesaurus were useful. These findings suggest that the thesaurus provided a useful source of search terms for the end-users involved in this study.

One of the issues in thesaurus-based query expansion research is to establish whether the additional terms suggested by thesauri are new to the user. In this study an analysis of users' initial terms and the thesaurus terms selected by users for query expansion was made. Results showed that the thesaurus provided new terms i.e. terms different from the users' initial terms for expanded queries. Users on average selected 2.2 new terms from the thesaurus for expanding their queries. This finding indicates that within the context of this study the thesaurus proved an effective tool assisting users in their term selection for query expansion.

To extent to which users have prior knowledge of the additional terms they find or select during their search process has been suggested to be of importance in understanding users' searching behaviour (Efthimiadis, 2000). This study found that in 50% of the cases where users found additional thesaurus terms they were not aware of those terms at the beginning of the search. This finding can be compared with those reported by Efthimiadis (2000) who found that around 44% of query expansion terms suggested by the system were terms representing new ideas. This finding suggests that thesauri have the potential to provide users with new search terms and to introduce different perspectives on a search topic, which users may not have been aware of or might not have thought of at the onset of the search. This finding concurs with Bates (1998) who stated that "the average person will recall (think up) only a fraction of the range of terms that are used to represent a concept or a name, but can take in a full screen of variants in an instant, and make a quick decision about desired terms for a given search". The finding that in half the cases users found new terms from the thesaurus is consistent with Jones and Hancock-Beaulieu (1994) who reported that the most frequently comment users made with respect to thesaurus navigation was that the thesaurus revealed new terms not previously thought of.

Users also provided three main reasons for choosing additional terms suggested by the thesaurus. These reasons were broadly categorised as follows: *narrowing down* (47%), *broadening* (22%) and *term relatedness* (18%). Comments made by the remaining 13% were: poor response to the first search, finding better terms, thesaurus terms matched original terms, provision of a richer vocabulary, confirming that the original search terms were correct. This finding reflects the fact that end-users are able to understand the underlying thesaurus relationships and to appreciate the ways in which they can broaden or narrow down their searches.

The search terms selected by academic staff and postgraduates were analysed to identify their relationships to users original terms. Academic staff chose more narrower and synonymous terms than did postgraduates who selected more broader and related terms. These differences represent patterns of thesaurus-based term selection for end-users with different levels of background knowledge. Academic staff have more extensive knowledge of their subject domain than postgraduates and this difference in background knowledge may have caused them to be more distinctive and specific in their selection of thesaurus terms.

While the particular focus of this study was on the search process, it also assessed search success by asking users to comment on the relevance of the retrieved result sets. It was found that result sets from around 63% of expanded searches were evaluated as being *very relevant* while the remaining 37% of result sets using expanded searches were judged to be *partially relevant*. This finding suggests that not only did the thesaurus aid the selection of search terms for query expansion but that it was also found to contribute to the retrieval of relevant records.

Another interesting finding concerning the retrieved records for all searches was the fact that while users evaluated 67% of searches as

being *very relevant*, they thought that in 54% of searches the results they viewed matched the topics they were interested in. This finding points to the fact that topics are not the only bibliographic surrogates on which users base their relevance judgement. It also points to the fact that while the retrieved results may not exactly match topics, they can still be very relevant to the user' information need. This difference between relevance evaluation and result-topic match assessment can also be viewed in the light of the various dimensions and manifestations of relevance judgement such as topical relevance and subjective relevance as indicated in the literature (Mizzaro, 1997; Ingwersen, 2001b).

6.8 Thesaurus Related Findings

Findings on users' search terms and their mapping to the CAB thesaurus suggest that over 80% of the search terms entered by the users were matched with the CAB thesaurus descriptors either exactly or partially. This finding can be compared with the findings reported by Fidel (1991b) who stated that in 80% of the cases that search specialists consulted thesauri, they found a match in the thesaurus. It is also interesting to compare the percentage of exact and partial matches reported in Fidel's study with those found in this investigation. She discovered that around 44% of the search terms entered by search specialists were exact matches while in this study 64% of search terms were exact match cases. This finding raises the issue of coverage and quality of thesauri in dealing with users' search terms. Fidel suggests that "among thesauri available today, a 'good' one provides at least an 80% match between request terms and descriptors". Therefore, having considered the fact that users in this study represented a relatively wide ranging subject area i.e. life sciences and in particular veterinary science, it can be suggested that the CAB thesaurus appears to have achieved an acceptable level of coverage with respect to users' terms.

The findings on different types of match between users' terms and the thesaurus also point to the notion of cognitive structures contributing to the IR interaction process as proposed by Ingwersen (1996). He suggests that there are different cognitive structures such as users' cognitive structure, the cognitive structures of thesaurus constructors and system designers, each of which has the potential to influence other cognitive structures. In this study the interaction between users and the thesaurus showed ways in which the cognitive structure inherent in the thesaurus affected users' cognitive structures in particular their selection of search terms.

Findings on the proportion of selected terms to browsed terms indicate that on average users selected over 14% of the terms they browsed. This finding can be contrasted with Jones et al. (1995) who found that during thesaurus navigation, on average users selected around 6% of the terms which they viewed. An explanation for this disparity may lie in experimental differences such as the thesaurus used, the retrieval system employed or the term ranking and interface options utilised.

One of the well-established debates in the area of controlled vocabularies has been the view that the role and power of free-text searching is not appreciated (Dubois, 1987; Fidel, 1991; Cousins, 1992; Rowley, 1994; Muddamalle, 1998). The Ovid search interface, which was utilised in the present investigation, provides a free text search facility within a thesaurus-based search setting whereby users can simultaneously choose both free text and thesaurus-based search. One interesting finding concerned the use of the *search as keyword* feature in the Ovid search interface. The feature helped users to choose their own search terms for free text searching in particular for statistical match cases where no exact or partial matches were found. This facility was useful in the sense that users did not have to change their search state and were able to add their term using this feature. The *search as keyword* facility

provided a combined search option taking advantage of both controlled vocabulary and free-text search. The finding that users made use of the *search as keyword* option in different types of match situations namely statistical, partial and exact match seems to indicate that the availability of the feature proved useful. In particular in cases of statistical match where users were provided with a large list of possibly relevant terms this was an often used option.

6.9 Perceptions of Users Towards Thesaurus Browsing/ Navigation Effect

The majority (93%) of users found thesaurus browsing and navigation to be easy. However, thesaurus browsing and navigation appeared to have different effects on academic staff when compared to postgraduates. While both groups agreed on certain common thesaurus effects such as narrowing down, broadening and provision of additional terms, there were differences in the ways in which the two groups perceived the function and use of the thesaurus. The comments made by postgraduates particularly focused on the broadening effect, the provision of new terms and terms of which they were not initially aware, as well as terms which were related to their area of interest. Academic staff, in contrast, tended to stress the narrowing down effect of the thesaurus, the provision of alternative terms, and an appreciation of the fact that topic familiarity plays a role in thesaurus browsing and navigation. This difference in perceptions was clearly found when comparing thesaurus terms selected by academic staff and postgraduates. Academic staff tended to choose more narrow and synonymous terms while postgraduates selected more broad and related terms. This difference may be explained by the fact that academic staff are more familiar with their subject area, have a richer vocabulary for query formulation and as a result tend to be more differentiating in browsing and term selection. This finding is consistent with the results from the empirical investigations carried out by Vakkari et

al. (2003) who found that an increase in subject domain knowledge generates a growth and differentiation of vocabulary reflected in the increased use of search terms, especially narrower terms.

Postgraduates encountered more difficulties while browsing the thesaurus than did academic staff. While postgraduates commented that they had to struggle to differentiate between Boolean operators, academic staff tended to comment on usability aspects of the interface such as the number of hits in front of each thesaurus terms or the visibility and size of buttons on the interface. An explanation for this difference could be the fact that academic staff were more experienced in using computers.

While both groups could recognise the semantic link between their own search terms and the thesaurus mapped terms, there were cases in non-exact match situations where users reported that the suggested search terms did not appear to be of relevance to the their search terms. The explanation for this finding lies in the fact that the Ovid thesaurus mapping facility uses a statistical algorithm based on some co-occurrence analysis of thesaurus descriptors. Therefore, users who enter a term such as *Ecoli 0157* will be shown the mapped term as well as terms such as *Australia* or *China* which have frequently co-occurred with the term *Ecoli 0157* in database records. Although this information could be useful per se, indicating where research on this disease has been carried out or where the disease is prevalent, users did not realise why these terms were displayed. This problem has also been articulated by Beaulieu et al. (1995) who stated that the lack of awareness regarding the source of terms may affect users' perceptions of potential query terms. There were also some thesaurus features which were found particularly helpful by users. These included the ability of the thesaurus in controlling British and American spelling to perform a comprehensive search, cross-referencing

between the acronyms and initials and their full forms, singular and plural forms of nouns and the vast coverage of animal species in the thesaurus. Participants from both groups mentioned that they had some degree of uncertainty as to what their search would yield when they selected either broader or narrower terms. Part of this problem could be due to the fact that the Ovid search interface does not provide instructions concerning the relationships between the users' terms and the thesaurus terms at the thesaurus mapping stage.

While the Boolean operators were explicitly available on the interface at both thesaurus mapping and thesaurus hierarchical states, participants from both groups mentioned that they required more help in notifying them that they could make use of or change the Boolean operators while selecting thesaurus terms for formulating their queries. Some users also mentioned that they were not sure about the difference between the AND and OR operators. This finding is in line with those reported by Vakkari (2000, 2003) who found that users needed help in formulating queries using the OR operator.

All subjects mentioned that the search interface and commands were easy to understand and use. There were certain interface features which users found particularly useful. These included: the e-mail facility whereby users could e-mail their search results, ease of use of the combine option and the presentation of terms and references. However, there were a number of usability problems pointed out by users which could be viewed as suggestions for improvement. These included: the size, position and labels of buttons; reducing the scrolling area in particular during thesaurus browsing; a help option for thesaurus terms and their usage; access to the results page; and more guidance on the use of Boolean operators while selecting thesaurus terms.

While in this experiment the Ovid advanced search mode was chosen in order to make use of its thesaurus mapping facility, the general impression given by the users suggests that they were able to take advantage of the interface features. This finding indicates that providing default search interface features may influence users' searching behaviour. This is in agreement with findings reported by Manglano et al. (1998) who suggested that automatically providing a mapping facility to the controlled vocabulary would be an advantage and that the selection of default options in an interface affects users' search behaviour and performance.

6.10 Summary

This chapter has discussed and suggested interpretations for the findings reported in chapters 4 and 5. The interpretations and explanations were organised on the following themes: user characteristics, topic complexity, topic familiarity, topic search experience, search term characteristics, interface usability, query reformulation, thesaurus-related findings and users' perception towards thesaurus browsing/navigation. The next chapter will present conclusions and discuss implications to be drawn from this discussion of findings together with recommendations for further research.

Chapter 7: Conclusions

7.1 Introduction

The overall aim of this research was to evaluate how and to what extent a thesaurus-enhanced search interface can assist end-users in selecting search terms for query expansion. Specifically, it aimed to ascertain users' attitude toward both the thesaurus and interface as tools for facilitating search term selection for query expansion. It also intended to identify searching and browsing behaviours of users interacting with a thesaurus-enhanced interface attached to a large bibliographic database. To achieve the objectives of this study the rationale, main research questions and hypotheses were presented (Chapter 1). Chapter 2 provided a background to and a justification for the present study by classifying and reviewing the relevant literature. The user-centred evaluation framework and experimental procedures, which were used to investigate user interaction with a thesaurus-enhanced search system, were examined in Chapter 3. Descriptive and qualitative findings were

dealt with in Chapter 4, while the results of quantitative analysis and hypothesis testing were provided in Chapter 5. Chapter 6 presented discussions and interpretations relating to the findings with reference to existing literature. This chapter serves two main purposes. First, it presents conclusions with respect to the five main research questions. Second, it discusses the limitations of the study and makes some suggestions for further research. The chapter and thesis end with some key final conclusions.

7.2 General Conclusions

The general conclusions provided in this section relate to the main research questions outlined in the Introduction.

- Are there common patterns of user behaviour in thesaurus-based browsing and searching?

The results of this research demonstrated behavioural differences between academic staff and postgraduates. A pattern that emerges from the cognitive and physical moves made suggests that within the context of a thesaurus-enhanced search environment, academic staff browse more thesaurus terms and make more selections than do postgraduates. This difference in cognitive behaviour reflects the fact that users with more subject knowledge, i.e. academic staff, tend to explore the thesaurus space more extensively and to be more cognitively involved in the examination of terms. A clear pattern of thesaurus term selection was also revealed. Academic staff chose more narrower and synonymous terms while postgraduates selected broader and related terms. This difference points to the fact that academics, with more extensive knowledge of their subject domain, tend to be distinctive and specific when selecting thesaurus terms when compared to postgraduates who tend to have a more limited understanding of the domain.

Another pattern of behaviour that emerged from the outcomes of this research was associated with gender. Female users selected fewer search terms, made less cognitive moves and viewed fewer records than did male users. Interesting differences were also observed in relation to search success, where female users tended to be significantly more satisfied with the number of terms they selected and with the results they retrieved.

- What relationships are there between users' initial query terms and the terms they select from the thesaurus for query expansion?

Users expanded 76% of their searches. The relationship between the query expansion terms selected from the thesaurus and the initial query terms were identified mostly as being hierarchical or associative. Narrower Terms constituted 31% of the total query expansion terms, Related Terms accounted for 28% and Broader Terms 10% of the expansion terms. 10% of query expansion terms were synonymous terms. The statistical relationship accounted for 22% of the expansion terms. These finding suggests that end-users are likely to select the majority of their expansion terms from either narrower or related terms.

The extent to which a thesaurus suggests new terms to users demonstrates the coverage and strength of the thesaurus in contributing to the selection of terms for query expansion. This study found that 50% of the additional search terms suggested by the thesaurus were identified as being terms of which the users were not aware or had not thought of using at the beginning of the search.

Users' comments on their reasons for selecting additional thesaurus terms also suggest that *narrowing down* was the most frequently noted reason for choosing those terms. This is consistent with the users' actual term selection behaviour as reported earlier. Thesaurus-based query

expansion also had an effect on the users' judgement of the relevance of the retrieved results. In 63% of expanded searches users evaluated the results as being *very relevant* compared to 37% which were evaluated as being *partially relevant*. These findings indicate that standard thesauri such as the one utilised in this research are capable of suggesting useful query expansion terms.

- Does topic complexity affect user-thesaurus interaction in general and search term selection in particular?

The study found that topic complexity did affect users' interaction with the thesaurus. Users searching for complex topics made significantly more cognitive and physical moves than those with simple topics. Topic complexity also had a significant effect on the number of terms selected. Users selected on average around four more terms in searching for complex topics than they did for simple topics. The query reformulation behaviour of users was affected by the degree of topic complexity. Complex topics caused users to make more query reformulations than did simple topics. Complex topic queries were expanded with twice as many terms as were simple topic queries. These findings suggest that topic complexity should be considered as a variable when analysing and studying users' search behaviour. Users who searched for complex topics evaluated the additional terms suggested by the thesaurus to be more useful than did those with simple topics. This suggests that complex topic searches may benefit more from the use of a thesaurus than do simple topic searches.

- Does topic familiarity affect user-thesaurus interaction?

The results of this study demonstrated that the extent to which a person is familiar with a topic affects their browsing behaviour in a thesaurus-enhanced search environment. Users who are moderately or very familiar

with their topics tended to browse more thesaurus terms than those who were not familiar. It was demonstrated that users in the *very familiar* and *moderately familiar* categories browsed twice as many thesaurus terms as did those in the *unfamiliar* category.

Users' prior topic search experience appeared to have a significant effect on their selection and evaluation of thesaurus terms. Users who had not previously searched for their topics found significantly more additional terms from the thesaurus than did those with prior topic search experience. It was also shown that users with no topic search experience found thesaurus terms more useful and closer to their original search topic than those who had previous topic search experience. These findings suggest that thesauri can be of particular benefit to users who are not familiar with a topic or have no prior experience of searching their topics. Another implication is that it may be possible to utilise a user's browse and selection behaviour to ascertain the extent to which they are familiar with a topic while searching.

- Does interface usability affect thesaurus browsing/navigation and other search behaviours?

The results suggest that interface usability is a factor affecting thesaurus browsing/navigation and other information searching behaviours. Users who evaluated the learning and using the interface as *very easy* also rated thesaurus browsing/navigation *very easy*. Usability was also found to have a relationship to the time users spent searching for their topic. Those who evaluated the ease of learning and using the interface as *very easy* spent significantly less time compared to those who rated it as *easy*.

Ease of thesaurus browsing/navigation was found to have significant relationships with overall number of physical moves, thesaurus terms browsed and query reformulation. Users who evaluated thesaurus

browsing/navigation as being *very easy* made fewer physical moves, browsed fewer thesaurus terms and made fewer query reformulations. These findings demonstrate that thesaurus interface usability should be viewed in the general context of users' information searching behaviour.

7.3 Particular Conclusions

7.3.1 Thesaurus Browsing and Interface Issues

The effect of thesaurus browsing and navigation as evaluated by users indicated two different perceptions with respect to the role and function of thesauri. Academic staff viewed the function of a thesaurus as being useful for narrowing down a search and providing alternative search terms. Postgraduates stressed the role of the thesaurus for broadening a search and providing new terms of which they had not been aware at the beginning of the search. This difference points to the fact that users with different levels of domain knowledge can take advantage of thesauri in different ways.

A general trend was observed with respect to the use of Boolean operators which suggested that the operators can be particularly useful within a thesaurus-enhanced search interface but that their implementation affects usage. As found in this study, all users made use of the AND operator and the majority (70%) of them introduced the OR operator in their search. However, they also noted that they required more help in order to make effective use of these operators.

7.3.2 Search Term and Search Result Issues

A significant relationship was found between search term satisfaction and the number of terms selected. It was shown that term satisfaction is likely to be associated with the selection of fewer search terms. The number of cognitive moves also influenced users' evaluation of term usefulness. Those who evaluated their terms to be *very useful* tended to make fewer

cognitive moves than did those who judged their terms to be *fairly useful*. These findings suggest that the quality of terms is more important than the quantity.

An association was found between users' relevance evaluation of result sets and the number of physical moves. Those searches whose result sets were evaluated as being *very relevant* were associated with significantly fewer physical moves. This finding reinforces the fact that quality of moves is more important than quantity. The analysis of relevance evaluation and topical match, i.e. match between users' original topic and the retrieved results, demonstrated that users judge these two criteria differently. While they judged 67% of searches as yielding *very relevant* results, they thought that in only 54% of searches did they find a *very good match*. This finding reflects the fact that relevance manifests itself in differing ways, seen here in terms of topical relevance and subjective relevance.

7.3.3 Thesaurus Term Issues

An analysis of the search states at which users selected their terms for query expansion showed that the thesaurus mapping state i.e. the state immediately after the initial search is conducted was found to contribute more to the query expansion process than did the hierarchical browsing state. Users were initially presented with a list of thesaurus terms to choose from and some users preferred to choose their terms from this mapping state while others investigated the context of individual thesaurus terms in the hierarchical state. This finding implies that the stages at which users are provided with the list of search terms as well as the richness and relevance of the suggested terms affect their term selection and query expansion. The findings on the users' search term selection at the thesaurus mapping and hierarchical states suggest that different users choose to select terms from either a list of terms or a thesaurus hierarchy. It was also shown that the number and quality of

search terms presented to the users and their personal preference for investigating the hierarchical context of a term are important issues which should be taken into consideration while studying search term selection of users.

The results for different types of match between users' terms and thesaurus descriptors demonstrated that over 80% of users' terms were matched exactly (64%) or partially (18%) to thesaurus descriptors and 18% of users terms matched statistically to the thesaurus. This finding reflects the strength and coverage of the CAB thesaurus in dealing with subject areas in the life sciences and veterinary medicine, the focus of users in this research. The finding that 28% of terms selected by users for query expansion were 'Related Terms' also indicates the strength of the associative relationships which exist within the CAB thesaurus. The finding that 20% of query expansion terms were identified as having statistical relationships with users' terms and 18% of users' initial terms were statistical matches point to the fact that thesauri have the potential to be enhanced through the use of statistical co-occurrence analysis techniques. This finding is particularly important because statistical matches were considered to be useful to their search by end-users.

7.4 Implications

The overall results of this research have implications for user education and the design of thesaurus-enhanced search interfaces.

A major finding of this study was that end-users with varying levels of subject knowledge and search experience were able to make use of a thesaurus-enhanced search interface for search term selection and query expansion. However, most users commented that they had not been introduced to thesauri as search tools before and pointed out that they would use the thesaurus for future searches. It is recommended that

online and database searching courses should incorporate training on thesaurus-based search options to improve user performance in conducting high quality searches.

It is suggested that the identification of major categories of move types and their relative frequencies will facilitate an evaluation of the contribution made by each type of move to the overall search process and can be of value in the design of IR interfaces. Results from this study indicate that complex topic searches require more query reformulation and involve users in viewing more search results. The integration of search term and search result spaces could be viewed as one mechanism for improving the interface to better support cognitive moves when performing complex topic searches. Another implication for interface design would be that facilities for query modification and expansion should be made more readily accessible to users carrying out searches associated with complex topics. Furthermore, a framework utilising cognitive and physical moves provides a method for assessing the extent to which users take advantage of browse, search and navigation features within an IR interface and can guide the interface design process such that cognitive and physical loads are reduced during the search process.

Results of this research demonstrated that users made more use of narrower terms and related terms than they did of other thesaurus terms. This finding has an implication for designing thesaurus interfaces in that it suggests that NT and RT terms should be presented at the top of the list of suggested terms. This design suggestion would be particularly useful for IR systems with term weighting and co-occurrence analysis techniques.

Statistical match cases played a major part in the process of suggesting alternative thesaurus terms to the user. This finding implies that statistical techniques such as the co-occurrence analysis can be used to test the usefulness and effectiveness of alternative term suggestion facilities based on thesauri.

The users selected their search terms for query expansion at both the thesaurus mapping and thesaurus hierarchical states. The first observation of this finding would be that the number and quality of search terms at different search states is an important issue to be considered in studying end-users' search term selection behaviour. The second observation lies in the fact that users' personal preferences play a role in selecting their terms from either the hierarchical state or the thesaurus mapping state. These two issues imply that the design of search interfaces should take into consideration user preferences with respect to term selection. From the design of the search interface perspective, there exists a trade-off between allowing users to browse a hierarchical structure or a flat list of search terms.

The level of users' topic familiarity and the associated search behaviour could be used to inform the design of personalised interaction with an IR system. For instance, users with a high degree of familiarity could be provided with additional features to browse more thesaurus terms.

The explicit availability of the AND and OR operators at both thesaurus mapping and thesaurus hierarchical states in the Ovid search interface promoted the use of these operators. However, users still require to be prompted to select either of them while combining thesaurus terms. It is suggested that user training in understanding Boolean logic will increase the correct and effective usage of operators and hence improve performance.

The list of thesaurus matched terms provided by the Ovid interface took up more than one screen in most cases. This demanded more physical effort from users to scroll up and down to view all terms. The list could be reduced in such a way that it takes up the standard screen and in this way users could easily view all terms without having to scroll up and down. A list of 13 to 15 terms will be a reasonable size to allow users to have a quick glimpse at all terms with no physical effort. This change would possibly have term ranking implication as to what thesaurus term should be displayed. This study found that users selected more Narrower and Related terms than the other thesaurus terms. The algorithm for term ranking could be designed in such a way as to favour Narrower and Related terms.

7.5 Limitations

Limitations of this study can be considered in three broad categories namely choice of subjects, experimental design, and the IR system used. The difficulties of involving users with real information needs in user studies are well documented. This study recruited volunteers from a university department. Although the number of subjects in both *status* groups, namely academic staff and postgraduates, was equal, it was not possible to control for factors such as age, gender or previous search experience. Therefore, the generalisability of the results of this study should be viewed in the light of this limitation.

Since the overall aim of this study required the involvement of real users with real search requests, the degree of control over the experimental design was constrained and as a result it was not possible to divide the subjects into control and experimental groups. The comparisons were mainly made between groups of users defined based on a range of variables rather than random effects or assigned search tasks.

This study utilised a commercial web-based information retrieval system to conduct the experiment. The system provided various thesaurus mapping features namely exact match, partial match and statistical match. Therefore, the search states defined to represent typical thesaurus-based search stages were restricted to those features. Different types of match also influenced the mapping of users' terms to the thesaurus descriptors. In the statistical match cases, some users were not certain about the statistically matched terms and could not understand where these terms came from. In light of this, statistical match cases were reported separately in order not to influence the general results, although they were also thesaurus terms.

7.6 Further Research

This study focused on end-user query expansion behaviour within the context of a thesaurus-enhanced search setting. A number of variables such as cognitive and physical moves, topic complexity, topic familiarity, and interface usability were examined. Further research should further investigate the nature and types of cognitive and physical moves in studying end-user search behaviour. It is suggested that cognitive and physical moves could be weighted based on their contribution to the whole search process and that this may provide a more sensitive method for evaluating users' search moves in relation to a particular task. When compared with search outcome measures such as relevance judgement or user satisfaction, cognitive and physical moves can be viewed as measures contributing to the evaluation of IR interaction and search efficiency. This approach to search moves also allows for an assessment of the extent to which an IR system involves users in cognitive or physical behaviour while searching for a topic.

User-centred query expansion research should examine different stages at which thesaurus-based query expansion is useful and the contribution which thesauri make to the whole search process. In this study the use of

a thesaurus for query expansion at the early stage of the search was found to be useful. Research is needed to study whether thesaurus-based query expansion at later stages would be as useful. Users in this study indicated a preference for hierarchical links when selecting search terms for query expansion. Further research is needed to investigate the behaviour of other end-user communities to establish whether this preference represents a domain-specific or even a thesaurus-specific characteristic. Further research should also explore the number of terms presented and the search features made available at each stage of the thesaurus-aided query expansion.

In this study topic complexity was defined based on the number of terms together with the Boolean operators utilised by users. This method of assessing topic complexity could be extended by incorporating a subjective evaluation by the user of topic complexity which would allow for a more balanced topic complexity measure. Simple and complex topics can have different query expansion implications both in terms of the expansion stage and other search behaviours. The finding that complex topics benefited more from thesaurus support than did simple topics could be tested in other domains using other types of thesaurus-enhanced IR systems.

Further research is needed in order to examine the effect of topic familiarity and subject knowledge on users' search behaviour in a thesaurus-aided search environment. To achieve a better understanding of users' knowledge of a topic, it is suggested that both topic familiarity and domain knowledge can be assessed and a single measure derived. The nature, type and relevance of search terms selected by users with various levels of knowledge and familiarity should also be investigated.

The results of this study showed that search interface usability and thesaurus browsing/navigation had relationships with other search

behaviour. Further research should address ways in which thesaurus interfaces can be made more usable in order to encourage users to take advantage of them. This calls for investigation into the search stages at which the thesaurus could be made available and the ways in which the use of thesauri for query expansion can be encouraged.

This study demonstrated that users with limited domain knowledge were able to take advantage of thesauri for selecting new terms and exploring the context of a particular term. It would be interesting to investigate the ways in which thesaurus terms might be useful to those new to a research area.

This research focused on veterinary science in an operational environment with end-users directly interacting with a thesaurus-enhanced search interface. It would be interesting to investigate the information searching behaviour of professionals from other closely related domains within the life sciences. There are several other thesaurus-enhanced search interfaces within operational IR systems in other subject domains. Further research is required to investigate thesaurus-based query expansion in user-centred environments using other domain-specific thesauri. This will shed light on users' query expansion behaviour and should show the extent to which this study's results are generalisable.

7.7 Conclusion of the study

This study investigated end-users' search term selection and query expansion behaviour in a thesaurus-enhanced search environment. It addressed issues surrounding users' cognitive and physical moves while interacting with such an environment.

The results indicated that thesauri are capable of assisting end-users in the selection of search terms for query formulation and expansion, in particular by providing new terms and ideas. This resulted in the majority of users expanding their queries and being successful in retrieving relevant information.

The categorisation of users' search topics as either simple or complex allowed for a detailed analysis of users' interaction with the thesaurus. It was shown that complex topic searches were subject to more query reformulations and also benefited more from the use of thesauri.

Over 80% of users' initial terms were matched with thesaurus terms and users selected the majority of their expansion terms from narrower and related terms. This points to a key role of thesauri within the search term selection process. It also reflects the fact that the quality and coverage of thesauri as well as their treatment of hierarchical and associative relationships are key elements in providing terms relevant to users' search terms.

A comparison of cognitive and physical behaviour between academic staff and postgraduates demonstrated that academic staff tend to browse more thesaurus terms. The findings on topic familiarity indicated that users with a high degree of topic familiarity also browse more thesaurus terms than do those who are not familiar with a topic.

The advantages of thesauri as perceived by academic staff and postgraduates reflected two different perspectives. Academic staff perceived thesauri as having a narrowing down effect on the search term selection process and as providing alternative search terms. In contrast, postgraduates stressed the broadening effect of thesauri and the provision of new terms and ideas. These two perceptions distinguish the

advantages that thesauri may have for end-users with varying degrees of subject knowledge.

This research has begun to identify elements that are important in evaluating user interaction with online thesauri. The use of cognitive and physical moves would appear to be an effective method for investigating end-user search behaviour and may ultimately contribute to the evaluation of IR interaction.

The findings of this study contribute to existing research that has examined end-user search behaviour in thesaurus-based search environments. The study has confirmed a number of results found in previous research. To the author's knowledge, it is the first study to explore variables such as topic complexity and topic familiarity in an operational thesaurus-enhanced search environment with end-users from disciplines other than library and information science. This study also adds to the research findings on user-centred evaluation of search term selection and thesaurus-based query expansion. The results from this study provide the basis for identifying ways in which thesaurus-enhanced interfaces can be designed and adapted in order to support users more adequately.

References

Agosti, M., Gradenigo, M.G., and Marchetti, P.G. (1991) Architecture and functions for a conceptual interface to very large online bibliographic collections. In*: Intelligent text and image handling, Proceedings of RIAO 91*. Barcelona, Spain, 2-5 April, Amsterdam, Netherlands: Elsevier, pp. 2-24.

Agosti, M. , Gradenigo, G., and Marchetti, P.G. (1992) A hypertext environment for interacting with large textual databases. *Information Processing and Management,* 28 (3), pp. 371-387.

Anderson, J. D. and Rowley, F. A. (1991) Building end-user thesauri from full-text. In Barbara H. Kwasink and Raya Fidel (eds.) *Advances in Classification Research, Proceedings of the 2nd ASIS SIG/CR Classification Research Workshop*. Medford, NJ: Learned Information, pp.1-13.

Bates, M.J. (1979) Information search tactics. *Journal of the American Society for Information Science*, 30 (4), pp. 205-214.

Bates, M.J. (1986) Subject access in online catalogs: a design model. *Journal of the American Society for Information Science,* 37 (6), pp. 357-376.

Bates, M.J. (1989) The design of browsing and berry picking techniques for online search interface. *Online Review*, 13 (5), pp. 407-431.

Bates, M.J. (1990a) Design for a subject search interface and online thesaurus for a very large record management database. In: Diane Henderson (ed.) *Proceedings of the 53rd Annual Meeting of the American Society for Information Science*, Toronto, Ontario, November 4-8 1990. Medford, NJ: Learned Information, pp.20-28.

Bates, M.J. (1990b) Where should the person stop and the information search interface start? *Information Processing and Management,* 26 (5), pp.575-591.

Bates, M.J. (1998) Indexing and access for digital libraries and the Internet: Human, database, and domain factors. *Journal of the American Society for Information Science*, 49 (13), pp.1185-1205.

Bates, M.J., Wilde, D.N., and Siegfried, S.L. (1993) An analysis of search terminology used by humanities searchers: The Getty online searching project report No 1. *Library Quarterly*, 63 (1), pp. 1-39.

Beaulieu, M. (1997) Experiments of interfaces to support query expansion. *Journal of Documentation*, 53 (1), pp. 8-19.

Beaulieu, M. (2000) Interaction in information searching and retrieval. *Journal of Documentation*, 56 (4), pp. 431-439.

Beaulieu, M., Fieldhouse, M., and Do, T. (1995) An evaluation of interactive query expansion in an online library catalogue with a graphical user interface. *Journal of Documentation*, 51 (3), pp. 225-243.

Beaulieu, M. and Robertson, S. (1996) Evaluating interactive systems in TREC. *Journal of the American Society for Information Science*, 47 (1), pp. 85-94.

Beaulieu, M., Do, T., Payne, A. and Jones, S. (1997) ENQUIRE Okapi Project British Library Research and Innovation Report 17. London: British Library.

Belkin, N.J. and Croft, B. (1992) Information filtering and information retrieval: two sides of the same coin? *Communication of the ACM*, 35 (12), pp. 29-38.

Belkin, N.J., Marchetti, P.G. and Cool, C. (1993a) BRAQUE: design of an interface to support user interaction in information retrieval. *Information Processing and Management,* 29 (3), pp. 325-344.

Belkin, N. J. (1993b) Interaction with texts: Information retrieval as information-seeking behaviour. *Information retrieval Von der Modellierung zur Anwendung*. Universitaetsverlag Konstanz, pp.55-66.

Belkin, N. J., Cool, C., Stein, A. and Thiel, U. (1995) Cases, scripts and information-seeking strategies: On the design of interactive information retrieval systems. *Expert Systems and Applications*, 9 (3), pp.379-395.

Belkin, N.J., Cool, C., Kelly, D., Lin, S., Park, S., Carballo, J. P. and Sikora, C. (2001) Iterative exploration, design and evaluation of support for query reformulation in interactive information retrieval. *Information Processing and Management*, 37 (3), pp. 403-434.

Bilal, D. and Kirby, J. (2002) Differences and similarities in information seeking: children and adults as Web users. *Information Processing and Management,* 38 (5), pp. 649-670.

Blocks, D., Binding, C., Cunliffe, D. and Tudhope, D. (2002) Qualitative evaluation of thesaurus-based retrieval. In: M. Agosti, C. Thanos (eds.) *Proceedings of 6th European Conference on Research and Advanced technology for Digital Libraries*. Rome, Italy, 16-18 September 2002. (Lecture Notes in Computer Science), Berlin: Springer, pp. 346-361.

Borgman, C.L., Case, D.O. and Meadow, C.T. (1989) The design and evaluation of a front-end user interface for energy researchers. *Journal of the American Society for Information Science,* 40 (2), pp. 99-109.

Borgman, C. L., Hirsch, S. G., and Walter, V. A., (1995) Children's searching behavior on browsing and keyword online catalogs: The science library catalog project. *Journal of the American Society for Information Science*, 46 (9), pp. 663-684.

Borgman, C.L., Hirsh, S.G., and Hiller, J. (1996) Rethinking online monitoring methods for information retrieval systems: From search product to search process. *Journal of the American Society for Information Science*, 47 (7), pp. 568-583.

Borland, P. and Ingwersen, P. (1997) The development of a method for the evaluation of interactive information retrieval systems. *Journal of Documentation*, 53 (3), pp. 225-250.

Brajnik, G., Mizzaro, S., and Tasso, C. (1996) Evaluating user interfaces to information retrieval systems: a case study on user support. *Proceedings of the 19th Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval.* August 18-22, 1996, Zurich, Switzerland, pp. 128-136.

Bystrom, K. and Jarvelin, K. (1995) Task complexity affects information seeking and use. *Information Processing and Management*, 31 (2), pp.191-213.

CAB Abstracts Database: http://www.cabi-publishing.org/Products/Database/Abstracts/Index.asp [last viewed 10 August 2003].

CAB Thesaurus: http://194.203.77.66/About.asp [last viewed 10 October 2003].

Cambridge Scientific Abstracts: http://www.csa.com [last viewed 16 January 2003].

Chen, H. and Dhar, V. (1991) Cognitive process as a basis for intelligent retrieval systems design. *Information Processing and Management*, 27 (7), pp. 405-432.

Chen, H., Schatz, B., Yim, T. and Fye, D. (1995) Automatic thesaurus generation for an electronic community system. *Journal of the American Society for Information Science,* 46 (3), pp. 175-193.

Chen, H. and Ng, T. (1995) An algorithmic approach to concept exploration in a large knowledge network (automatic thesaurus

consultation): Symbolic Branch-and-Bound search vs. connectionist Hopfield net activation. *Journal of the American Society for Information Science*, 46 (5), pp. 348-369.

Chen, H. and Ng, T. D. (1997) A concept space approach to addressing the vocabulary problem in scientific information retrieval: An experiment on the worm community system. *Journal of the American Society for Information Science*, 48 (1), pp. 17-31.

Chen, H., Martinez, J., Kirchhoff, T., Ns, D. and Schatz, B. R. (1998) Alleviating search uncertainty through concept associations: Automatic indexing, co-occurrence analysis, and parallel computing. *Journal of the American Society for Information Science*, 49 (3), pp. 206-216.

Cochrane, P.A. (1992) Indexing and searching thesauri, the Janus or Proteus of information retrieval. In: Williamson, N.J., Hudon, M. (ed.) *Classification Research for Knowledge Organization*, FID, pp. 161-178.

Cousins, S.A. (1992) Enhancing access to OPACs: Controlled vs natural language. *Journal of Documentation*, 48 (3), pp. 291-309.

Ding, Y., Ghowdhury, G.G., and Foo, S. (2000) Incorporating the results of co-word analyses to increase search variety for information retrieval. *Journal of Information Science*, 26 (6), pp. 429-451.

Dubois, C.P.R. (1987) Free text vs controlled vocabulary: A reassessment. *Online Review*, 11 (4), pp. 243-253.

Efthimiadis, E. N. (1993) A user-centred evaluation of ranking algorithms for interactive query expansion. Korfhage et.al. (eds.) *Proceedings of the 16th Annual International Conference on Research and Development in Information Retrieval of the Association for Computing Machinery Special Interest Group on Information Retrieval (ACM/SIGIR)*. New York: ACM, pp. 146-159.

Efthimiadis, E.N. (1995) User choices: A new yardstick for the evaluation of ranking algorithms for interactive query expansion. *Information Processing and Management*, 31 (4), pp. 605-620.

Efthimiadis, E.N. (1996) Query expansion. In: *Annual Review of Information Science and Technology*, edited by M.E. Williams. Medford, NJ. : Information Today, pp. 121-187.

Efthimiadis, E.N. (2000) Interactive query expansion: A user-based evaluation in a relevance feedback environment. *Journal of the American Society for Information Science*, 51 (11), pp. 989-1003.

Ellis, D. (1989) A behavioural approach to information retrieval system design. *Journal of Documentation*, 45 (3), pp.171-212.

ERIC Search Wizard: http://www.ericae.net/scripts/ewiz/home0.htm [last viewed 20 January 2003].

Fenichel, C.H. (1981) Online searching: Measures that discriminate among users with different types of experiences. *Journal of the American Society for Information Science*, 32 (1), pp. 23-32.

Fidel, R. (1984) Online searching styles: a case-study-based model of searching behaviour. *Journal of the American Society for Information Science*, 35 (4), pp. 211-221.

Fidel, R. (1985) Moves in online searching. *Online Review*, 9 (1), pp. 61-74.

Fidel, R. (1986) Towards expert systems for selection of search keys. *Journal of the American Society for Information Science*, 37 (1), pp. 37-44.

Fidel, R. (1991a) Searcher's selection of search keys: I. The selection routine. *Journal of the American Society for Information Science*, 42 (7), pp. 490-500.

Fidel, R. (1991b) Searchers' selection of search keys: II. Controlled vocabulary or free-text searching. *Journal of the American Society for Information Science*, 42 (7), pp. 501-514.

Fidel, R. (1991c) Searchers' selection of search keys: III. Searching styles. *Journal of the American Society for Information Science*, 42 (7), pp. 515-527.

Fidel, R. (1992) Who needs controlled vocabulary? *Special Libraries*, 83 (1), pp. 1-9.

Ford, N., Miller, D. (1996) Gender differences in Internet perception and use. *Aslib Proceedings*, 48 (7/8), pp. 183-192.

Ford, N., Miller, D., Moss, N. (2001) The role of individual differences in Internet searching: an empirical study. *Journal of the American Society for Information Science and Technology*, 52 (12), pp. 1049-1066.

Fowkes, H. and Beaulieu, M. (2000) Interactive searching behaviour: Okapi experiment for TREC-8. *The BCS/ IRSG 22nd Annual Colloquium on Information Retrieval Research*, Cambridge, 5-7 April, Cambridge: BCS/IRSG, pp. 47-56.

Furnas, G.W., Landauer, T. K., Gomez, L.M. and Dumais, S.T. (1987) The vocabulary problem in human-system communication, *Communications of ACM,* 30 (11), pp. 964-971.

Ganzmann, J. (1990) Criteria for the evaluation of thesaurus software. *International Classification,* 17 (3-4), pp. 148-157.

Gordon, A. and Domeshek, E.A. (1998) Déjà vu: a knowledge-rich interface for retrieval in digital libraries. In: *Proceedings of the 1998 International Conference on Intelligent User Interfaces (IUI),* January 6-8 1998, San Francisco. New York: ACM Press, pp.127-134.

Greenberg, J. (2001a) Automatic query expansion via lexical-semantic relationships. *Journal of the American Society for Information Science and Technology*, 52 (5), pp. 402-415.

Greenberg, J. (2001b) Optimal query expansion (QE) processing methods with semantically encoded structures thesauri terminology. *Journal of the American Society for Information Science and Technology*, 52 (6), pp. 487-498.

Guidelines for TREC- 2003 Interactive Track: http://www.ted.cmis.csiro.au/TRECInt/guidelines.html [last viewed 26 November 2003].

Hansen, P. (1999) User interface design for IR interaction: A task-oriented approach. In: Aparac, T., T. Saracevic, P. Ingwersen and P. Vakkari (Eds.). CoLIS 3. Proceedings of the *Third International Conference on the Conceptions of the Library and Information Science*, Dubrovnik, Croatia, 23-26 May 1999, pp. 191-205.

Harman, D. K. (1988) Toward interactive query expansion. Chiaramella, Yves (ed.) *Proceedings of the Association for Computing Machinery Special Interest group on Information Retrieval (ACM/SIGIR) 11th International Conference on Research and Development in Information Retrieval*. Grenoble,: Presses Universitaires de Grenoble, pp. 321-331.

Harter, S. and Hert, C.A. (1998) Evaluation of information retrieval systems: Approaches, issues and methods. In Martha E. Williams (ed.), *Annual Review of Information Science and Technology*, volume 32, Washington, D.C: American Society for Information Science,1997, pp. 3-94.

Hearst, M. A. (1999) User interfaces and visualisation. In: Baeza-Yates, R. and Ribeiro-Neto, B. (eds.), *Modern Information Retrieval*. New York: Addison Wesley, pp. 257-323.

Hearst, M.A. and Karadi, C. (1997) Cat-a-Cone: An interactive interface for specifying searches and viewing retrieval results using a large category hierarchy. In: *Proceedings of the 20th Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval of the Association for Computing Machinery Special Interest*

Group on Information Retrieval. Philadelphia, PA, USA, July 27- 31, 1997. New York: ACM, pp. 246-255.

Howard, H. (1982) Measures that discriminate among online users with different training and experience. *Online Review*, 6 (4), pp. 315-326.

Hsieh-Yee, I. (1993) Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science*, 44 (3), pp. 161-174.

Hsieh-Yee, I. (2001) Research on Web search behaviour. *Library and Information Science Research*, 23 (2), pp 167-185.

Iivonen, M. (1995) Searchers and searchers: Differences between the most and least consistent searchers. Fox, E., Ingwersen, P., and Fidel, R. (eds.) *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval of the Association for Computing Machinery Special Interest Group on Information Retrieval*. New York: ACM, pp. 149-157.

Iivonen, M. and Sonnenwald, D.H. (1998) From translation to navigation of different discourses: A model of search term selection during pre-online stage of the search process. *Journal of the American Society for Information Science*, 49 (4), pp. 312-326.

Ingwersen, P. (1982) Search procedures in the library analysed from cognitive point of view. *Journal of Documentation*, 38 (3), pp. 165-191.

Ingwersen, P. (1992) *Information retrieval interaction*. London: Taylor Graham Publishing.

Ingwersen, P. (1996) Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *Journal of Documentation*, 52 (1), pp. 3-50.

Ingwersen, P. (2001a) Cognitive information retrieval. In: Martha E. Williams (ed.) *Annual Review of Information Science and Technology (ARIST)*, Vol. 34, ASIS, pp. 3-52.

Ingwersen, P. (2001b) Users in context. In: Agosti, M., Crestani, F. and Pasi, G. (eds.) *Lectures on Information Retrieval.* Bonn: Springer Verlag, 2001, (Lecture Notes in Computer Science: 1980), pp.157-178.

Jansen, B. J. (2000) The effect of query complexity on Web searching results. *Information Research*, 6 (1) Available at: http://InformationR.net/ir/6-1/paper87.html. [last viewed 26 March 2003].

Jansen, B. J., Spink, A., and Saracevic, T. (2000) Real life, real users, Real needs: A study and analysis of user queries on the web. *Information Processing and Management*, 36 (2), pp. 207-227.

Jing, Y. and Croft, W. B. (1994) The association thesaurus for information retrieval. Proceedings of RIAO '94: Intelligent Multimedia Information Retrieval Systems and Management. Paris: CID, pp. 146-160.

Johnson, E.H. (1997) Using IODyne: Illustrations and examples. In: Cochrane, P.A. and Johnson, E. (eds.) *Visualizing Subject Access for 21st* Century Information Resources: Proceedings of the 34th Annual Clinic on *Library Applications of Data Processing*. Graduate School of Library and Information Science, University of Illinois, Illinois, March 1997, Champaign, Ill.: University of Illinois at Champaign-Urbana, School of Library and Information Science, pp. 80-93.

Johnson, E.H. and Cochrane, P.A. (1995) A hypertextual interface for a searcher's thesaurus. In: *Digital Libraries '95: Proceedings of the Second Annual Conference on the Theory and Practice of Digital Libraries*. Austin, Texas June 11-13 1995.

Joho, H., Coverson, C., Sanderson, M. and Hancock-Beaulieu, M. (2002) Hierarchical presentation of expansion terms. In: *Proceedings of 2002 ACM Symposium on Applied Computing*. Universidad Carlos III De Madrid, Madrid, Spain, March 11-14, 2002. New York: ACM Press, 2002, pp 645-649.

Jones, S. and Hancock-Beaulieu, M. (1994) Support strategies for interactive thesaurus navigation. In: Albrechtsen, H. & Ørnager, S.(eds.) *Knowledge organization and quality management: proceedings of the Third International ISKO Conference, 20-24 June 1994, Copenhagen, Denmark*. Advances in Knowledge Organization, 4. Frankfurt/Main: Indeks, pp. 366-373.

Jones, S., Gatford, M., Hancock-Beaulieu, M., Robertson, S.E. , Walker, W. and Secker, J. (1995) Interactive thesaurus navigation: Intelligence rules Ok? *Journal of the American Society for Information Science,* 46 (1), pp 52-59.

Jones, S. , Gatford, M., Do, T. and Walker, A. (1997) Transaction logging. *Journal of Documentation*, 53 (1), pp. 35-50.

Kelly, D. and Cool, C. (2002) The effects of topic familiarity on information search behaviour. *Proceedings of Joint Conference on Digital Libraries (JCDL),* July 13-17 2002, Portland, Oregon. New York: ACM Press, 2002, pp. 74-75.

Kekäläinen, J. and Jarvelin, K. (1998) The impact of query structure and query expansion on retrieval performance. Croft, W. B. et al. (eds.) Proceedings of the Association for Computing Machinery Special Interest Group on Information Retrieval (ACM/SIGIR) 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. 98. Melbourne, New York, NY: ACM Press, pp. 130-137.

Knapp, S.D., Cohen, L.B., and Judes, D.R. (1998) A natural language thesaurus for humanities. *Library Quarterly*, 68 (4), pp. 406-430.

Koenemann, J., Quatrain, R. and Belkin, N. (1994) New tools and old habits: The interactive searching behavior of expert online searchers using INQUERY In TREC-3. Harman, D. (ed.) *Proceedings. of the Third Text REtrieval Conference*, Washington D.C.,1994, GPO, pp. 144-177.

Korn, F. and Shneiderman, B. (1995) *Navigating terminology hierarchies to access a digital library of medical images*. University of Maryland Technical Report HCIL-TR-94-03.

Kristensen, J. (1993) Expanding end-users' query statements for free text searching with a search-aid thesaurus. *Information Processing and Management*, 29 (6), pp. 733-744.

Kristensen, J. and Jarvelin, K. (1990) The effectiveness of a searching thesaurus in free text searching of a full-text database. *International Classification*, 17 (2), pp. 77-84.

Kulthau, C.C. (1991) Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42 (5), pp. 361-371.

Lin, X. (1999) Visual MeSH. In: Hearst, M. et al.(eds.) *Proceedings of 22nd Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval of the Association for Computing Machinery Special Interest Group on Information Retrieval,* Berkeley, CA, USA, August 15-19 1999, New York: ACM, pp 317-318.

LISAnet: http://www.lisanet.co.uk/content/free/welcome.asp [last viewed March 2002].

Lopez-Huertas, M. J. (1997) Thesaurus structure design: A conceptual approach for improved interaction. *Journal of Documentation*, 53 (2), pp.139-177.

Mandala, R., Tokunaga, T. and Tanaka, H. (1999) Combining multiple evidence from different types of thesaurus for query expansion. *Proceedings of the Association for Computing Machinery Special Interest Group on Information Retrieval (ACM/SIGIR) 22th Annual International*

Conference on Research and Development in Information Retrieval. August 15-19, Berkeley, 1999. New York: ACM press, 1999, pp. 191-197.

Magennis, M. and Van Rijsbergen, C. J. (1997) The potential and actual effectiveness of interactive query expansion. *Proceedings of the Association for Computing Machinery Special Interest Group on Information Retrieval (ACM/SIGIR) 20th Annual International Conference on Research and Development in Information Retrieval.* 27-31 July, Philadelphia, 1997. New York, NY:ACM Press, 1997, pp. 324-332.

Majid, S., Eisenschitz, T.S., and Anwar, M.A. (1999) Library use pattern of Malaysian agricultural scientists. *Libri*, 49 (4), pp. 225-235.

Majid, S., Anwar, M.A., Eisenschitz, T.S. (2000) Information needs and information seeking behaviour of agricultural scientists in Malaysia. *Library and Information Science Research* 22 (2), pp. 145-163.

Manglano, V., Beaulieu, M. and Robertson, S. (1998) Evaluation of interfaces for IRS: Modelling end-user searching behaviour. *20th British Computing Society Information Retrieval Special Group (BCS/IRSG) Colloquium on Information Retrieval*, 25-27 March 98, Grenoble (France), (Electronic Workshops in Computing), Grenoble: Springer-Verlag, 1998, pp. 137-146.

Marchionini, G. (1989) Information-seeking strategies of novices using a full-text electronic encyclopaedia. *Journal of the American Society for Information Science*, 40 (1), pp. 54-66.

Marchionini, G., Meadow, C., Dwiggins, S., Lin, X., Wang, J. and Yuan, W. (1991) A study of user interaction with information retrieval interfaces: progress report. *The Canadian Journal of Information Science*, 16 (4), pp. 42-59.

Marchionini, G. (1992) Interfaces for end-user information seeking. *Journal of the American Society for Information Science* 43 (2), pp.156-163.

Marchionini, G. and Komlodi, A. (1998) Design of interfaces for information seeking. In: Williams, Martha E. (ed.), *Annual Review of Information Science and Technology* volume 33. Medford, NJ: Learned Information, pp. 89-130.

Mandala, R., Tokunaga, T. and Tanaka, H. (1999) Combining multiple evidence from different types of thesaurus for query expansion. *Proceedings of the Association for Computing Machinery Special Interest Group on Information Retrieval (ACM/SIGIR) 22nd Annual International Conference on Research and Development in Information Retrieval.* New York : ACM, pp. 191-197.

Markey, K., Atherton, P., and Newton, C. (1980) An analysis of controlled and free-text search statements in online searches. *Online Review*, 4 (3), pp. 225-236.

McMath, C.F., Tamaru, R.S. and Rada, R. (1989) A graphical thesaurus-based information retrieval system. *International Journal of Man-Machine Studies,* 31 (2), pp. 121-147.

Meadow, C., Wang, J., and Yuan, W. (1995) A study of user performance and attitudes with information retrieval interfaces. *Journal of the American Society for Information Science*, 46 (7), pp. 490-505.

Milstead, J. L. (1991) Specifications for thesaurus software. *Information Processing and Management,* 27 (2-3), pp. 165-175.

Milstead, J. L. (1997) Use of thesauri in the full-text environment. In: Cochrane, P.A. and Johnson, E.H. (eds.) *Visualizing Subject Access for 21st Century Information Resources: Proceedings of the 34th Annual Clinic on Library Applications of Data Processing*, Champaign, IL: Graduate School of Library and Information Science, University of Illinois, 1997, pp. 28-38.

Mizzaro, S. (1997) Relevance: The whole history. *Journal of American Society for Information Science*, 48 (9), pp. 810-832.

Muddamalle, M.R. (1998) Natural language versus controlled vocabulary in information retrieval: A case study in soil mechanics. *Journal of the American Society for Information Science*, 49 (10), pp. 81-887.

Nielsen, M.L. (2001) A framework for work task-based thesaurus design. *Journal of Documentation*, 57 (6), pp. 774-797.

Nweke, K.M. (1995) Information methods of human and veterinary medical scientists (HVMS) in Borno state, Nigeria. *Library and information Science Research*, 17 (1), pp. 41-48.

Ovid Technologies: http://www.ovid.com/ [last viewed September 2003].

Park, S. (2000) Usability, user preferences, effectiveness and user behaviours when searching individual and integrated full-text databases: Implications for digital libraries. *Journal of the American Society for Information Science and Technology*, 51 (5), pp. 456-468.

Patton, M. Q. (2002) *Qualitative research and evaluation methods*. 3rd ed. London :Sage Publications Inc.

Penniman, W. D. and Dominick, W. D. (1980) Monitoring and evaluation of on-line information system usage. *Information Processing and Management,* 16 (1), pp. 17-35.

Perez, E. (1982) Text enhancement: controlled vocabulary vs. Free text. *Special Libraries*, 73 (July), pp.183-192.

Piternick, A. (1984) Searching vocabularies: A developing category of online searching tools. *Online Review*, 8 (5), pp. 441-449.

Pollard, R. (1993) A hypertext-based thesaurus as subject browsing aid for bibliographic database, *Information Processing and Management,* 29 (3), pp. 345-357.

Pollitt, S. (1987) CANSEARCH: An expert system approach to document retrieval. *Information Processing and Management,* 23 (2), pp.119-138.

Pollitt, A.S. (1988) Common query interface using MenUSE- a menu-based search engine. In: *Online Information Meeting 1988: Proceedings of the 12th International Online Meeting*, London, 6-8 December 1988. Oxford: Learned Information, pp. 445-457.

Pollitt, A.S., Ellis, G. P., Smith, M. P., Gregory, M. R., Li, C. S. and Zangenberg, H. (1993) A common query interface for multilingual document retrieval from databases of the European Community institutions. In: *Online Information 1993: Proceedings of the 17th International Online Information Meeting*, London, 7-9 December 1993. Oxford: Learned Information, pp. 47-61.

Pollitt, A.S., Ellis, G.P. and Smith, M.P. (1994a) HIBROWSE for bibliographic databases. *Journal of Information Science* 20 (6), pp. 413-426.

Pollitt, A. S., Ellis, G.P. and Smith, M.P. (1994b) Improving search quality using thesauri for query specification and the presentation of search results. In: *Knowledge Organization and Quality Management: Proceedings of the 3rd International ISKO Conference*, Copenhagen, 20-14 June 1994. Frankfurt, Germany: INDEK Verlag, pp. 382-389.

Pollitt, A.S., Smith, M.P. and Braekevelt, P.A.J. (1996a) View-based searching systems- a new paradigm for information retrieval based on faceted classification and indexing using constraining knowledge-based views. In: Johnson, C. and Dunlop, M. (eds.), *Information Retrieval and Human-Computer Interaction 1996: Proceedings of the joint workshop of IR and HCI Special Interest groups of the British Computing Society*, Glasgow University, 17 September 1996, Glasgow: Glasgow University, 1996, pp. 73-77.

Pollitt, A. S., Treglown, M., and Braekevelt, P.A.J. (1996b) View-based searching systems--progress towards effective disintermediation. In: *Online Information 1996: Proceedings of the 20th International Online Meeting*, London, England, December 1996. Oxford: Learned Information Europe Ltd., pp. 433-446.

Pollitt, A.S. (1997) The key role of classification and indexing in view-based searching. In: *IFLA '97: Proceedings of the 63rd IFLA General* Copenhagen, Denmark, August 31-September 5 1997, Booklet 4, Section on Classification and Indexing Session 95 Paper 009-CLASS-1-E.

PubMed: http://www4.ncbi.nlm.nih.gov/PubMed/ [last viewed 15 March 2002].

Qiu, L. (1993) Analytical Searching vs. Browsing in Hypertext Information Retrieval Systems. *Canadian Journal of Information and Library Science,* 18 (4), pp. 1-13.

Qiu, Y. and Frei, H. P. (1993) Concept-based query expansion. *Proceedings of the Association for computing Machinery special Interest Group on Information Retrieval (ACM/SIGIR) 16th International ACM SIGIR Conference on Research and Development in Information Retrieval.* 27 June- 1 July, Pittsburgh, PA. New York, NY: ACM Press, pp. 160-169.

Robertson, S.E. (1990) On term selection for query expansion. *Journal of Documentation*, 46 (4), pp. 359-364.

Robertson, S. E. and Hancock-Beaulieu, M. (1992) On the evaluation of IR systems. *Information Processing and Management*, 28 (4), pp. 457-466.

Robertson, S. E. and Beaulieu, M. (1997) Research and evaluation in information retrieval. *Journal of Documentation*, 53 (1), pp. 51-57.

Robertson, S. E., Walker, S., and Beaulieu, M. (1997) Laboratory experiments with Okapi: Participation in the TREC programme. *Journal of Documentation*, 53 (1), pp. 20-32.

Rowley, J. (1994) The controlled versus natural indexing languages debate revisited: A perspective on information retrieval practice and research. *Journal of Information Science*, 20 (2), pp. 108-119.

Salton, G. and Buckley, C. (1988) Term-weighting approaches in automatic text retrieval. *Information Processing and Management*, 24 (5), pp. 513-523.

Saracevic, T., Kantor, P., Chamis, A. Y. and Trivison, D. (1988) A study of information seeking and retrieving. I. Background and methodology. *Journal of the American Society for Information Science*, 39 (3), pp. 161-176.

Saracevic, T. and Kantor, P. (1988) A study of information seeking and retrieving: Users, questions and effectiveness. *Journal of the American Society for Information Science*, 39 (3), pp. 177-196.

Saracevic, T. and Kantor, P. (1988) A study of information seeking and retrieving .III. Searchers, searches, and overlap. *Journal of the American Society for Information Science*, 39 (3), pp. 197-216.

Saracevic, T. (1996) Relevance reconsidered. Information science: Integration in perspectives. *Proceedings of the Second Conference on Conceptions of Library and Information Science*. Copenhagen: The Royal School of Librarianship, pp. 201-218.

Saracevic, T. (1995) Evaluation of evaluation in information retrieval. In: Fox, E. A., P. Ingwersen, and F. Fidel (eds.). *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'95)*, Seattle, Washington, USA, July 9-13, 1995. New York: ACM, 1995, pp.138-146.

Saracevic, T. (1997) The stratified model of information retrieval interaction: Extension and applications. In Schwartz, C., Rorvig, M. (ed.) *ASIS'97 : Proceedings of the American Society for Information Science. Vol. 34*. Silver Spring, Maryland: ASIS, 1997, pp. 313-327.

Saracevic, T. (1999) Information Science. *Journal of the American Society for Information Science*, 50 (12), pp. 1051-1063.

Savage-Knepshield, P.A. and Belkin, N.J. (1999) Interaction in information retrieval: trends over time, *Journal of the American Society for Information Science* 50 (12), pp. 1067-1082.

Schatz, B. R., Johnson, E.H. and Cochrane, P.A. (1996) Interactive term suggestion for users of digital libraries: using subject thesauri and co-occurrence lists for information retrieval. *Proceedings of the 1st Association for Computing Machinery International Conference on Digital libraries*. Bethesda, MD: ACM Press, pp.126-133.

Schutze, H. and Pedersen, J. (1997) A co-occurrence-based thesaurus and two applications to information retrieval. *Information Processing and Management*, 33 (3), pp. 307-318.

Shapiro, C.D. and Yan, P.F. (1996) Generous tools: Thesauri in digital libraries. In: *National Online Meeting: Proceedings of 17th National*

Online Meeting. New York, May 14-16 1996 Medford, New Jersey: Information Today Inc., pp. 323-332.

Shiri, A.A. and Revie, C. (2000) Thesauri on the web: Current developments and trends. *Online Information Review*, 24 (4), pp. 273-279.

Shiri, A. A. and Revie, C. (2001) User - thesaurus interaction in a web-based database: an evaluation of users' search term selection behaviour. In: *Proceedings of the Infotech Oulu International workshop on Information Retrieval*, Oulu, Finland 19-21 September 2001. Oulu: University of Oulu, pp. 23-32.

Shiri, A. A., Revie, C. and Chowdhury, G. (2002a) Thesaurus-Assisted Search Term Selection and Query Expansion: A Review of User-Centred Studies. *Knowledge Organization*, vol. 29 (1), pp.1-19.

Shiri, A. A., Revie, C. and Chowdhury, G. (2002b) Thesaurus-enhanced search interfaces. *Journal of Information Science*, 28 (2), pp.111-122.

Shiri, A. A., Revie, C. and Chowdhury, G. (2002c) Assessing the Impact of User Interaction with Thesaurual Knowledge Structures: A Quantitative Analysis Framework. In: C*hallenges in knowledge representation and organization for the 21st century: integration of knowledge across boundaries: Proceedings of the 7th Internationl Society for Knowledge Organization (ISKO) Conference*, Granada, Spain, 10-13 July 2002. Wurzburg: Ergon Verlag, pp. 493-499.

Shneiderman, B. (1998) Designing the User Interface: Strategies for Effective Human-Computer Interaction. Reading, MA: Addison-Wesley Publishing.

Siegfried, S., Bates, M.J., and Wilde, D.N. (1993) A profile of end-user searching behaviour by humanities scholar: The Getty online Searching project report No 2. *Journal of the American Society for Information Science*, 44 (5), pp. 273-291.

SilverPlatter: http://www.silverplatter.com [last viewed March 2002].

Solomon, P. (1993) Children's information retrieval behaviour: A case analysis of an OPAC. *Journal of the American Society for Information Science*, 44 (5), pp. 245-264.

Spark Jones, K. (1979) Search term relevance weighting given little relevance information. *Journal of Documentation*, 35 (1), pp. 30-48.

Spink, A. (1994a) Term relevance feedback and query expansion: Relation to design. Croft, W. B. and Van Rijsbergen, C. J. (eds.) *Proceedings of the Association for Computing Machinery Special Interest Group on Information Retrieval (ACM/SIGIR) 17th Annual International Conference on Research and Development in Information Retrieval.* Berlin: Springer-Verlag, pp. 81-90.

Spink, A. (1994b) Term relevance feedback and mediated database searching: Implications for information retrieval practice and systems design. *Information Processing and Management*, 31 (2), pp.161-171.

Spink, A. (1996) Multiple search sessions model of end-user behaviour: An exploratory study. *Journal of the American Society for Information Science*, 47 (8), pp. 603-609.

Spink, A., Goodrum, A., Robins, D. and Wu, M. (1996) Elicitations during information retrieval: implications for IR system design. *Proceedings of the Association for Computing Machinery Special Interest Group on Information Retrieval (ACM/SIGIR) 19th Annual International Conference on Research and Development in Information Retrieval.* August 18-22, 1996, Zurich. New York, NY: ACM Press, pp.120-127.

Spink, A. and Saracevic, T. (1997) Interaction in information retrieval: Selection and effectiveness of search terms. *Journal of the American Society for Information Science*, 48 (8), pp. 741-761.

Spink, A., Bateman, J., and Jansen, B.J. (1999) Searching the web: A survey of Excite users. *Internet Research: Electronic Networking Applications and Policy*, 9 (2), pp. 117-128.

Spink, A., Jansen, B.J., and Ozmultu, H.C. (2000) Use of query reformulation and relevance feedback by Excite users. *Internet Research: Electronic Networking Applications and Policy*, 10 (4), pp. 317-328.

Spink, A., Wolfram, D., Jansen, B.J. and Saracevic, T. (2001) Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology*, 52 (3), pp. 226-234.

Spink, A. (2002) A user-centered approach to evaluating human interaction with Web search engines: An exploratory study. *Information Processing and Management*, 38 (3), pp. 401-426.

Spink, A., Wilson, T. D., Ford, N., Foster, A. and Ellis, D. (2002) Information seeking and mediated searching study. Part 3. Successive searching. *Journal of the American Society for Information Science and Technology*, 53 (9), pp. 716-727.

Sugar, W. (1995) User-centered perspective of information retrieval research and analysis methods. In: Williams, M. E (ed.). *Annual Review of Information Science and Technology (ARIST)*, Vol. 30, Medford, N.J.: ASIS, 1995, pp. 77-109.

Sutcliffe, A.G., Ennis, M., and Watkinson, S.J. (2000a) Empirical studies of end-user information searching. *Journal of the American Society for Information Science*, 51 (13), pp. 1211-1231.

Sutcliffe, A.G., Ennis, M., Hu, J. (2000b) Evaluating the effectiveness of visual user interfaces for information retrieval. *International journal of Human-Computer Studies*, 53 (5), pp. 741-763.

Svenonius, E. (1986) Unanswered questions in the design of controlled vocabularies. *Journal of the American Society for Information Science*, 37 (5), pp. 331-340.

Tague-Sutcliffe, J. (1992) The pragmatics of information retrieval experimentation, revisited. *Information Processing and Management,* 28 (4), pp. 467-490.

Text REtrieval Conference (TREC) proceedings: http://trec.nist.gov [last viewed April 2003].

Van Rijsbergen, C.J., Harper, D. J., and Porter, M.F. (1981) The selection of good search terms. *Information Processing and Management*, 17 (2), pp. 77-91.

Vakkari, P. (2000) Cognition and changes of search terms and tactics during task performance: A longitudinal case study. *Proceedings of RIAO 2000, Content-based Multimedia Information Access RIAO Conference, Paris, April 12-14 2000.* Paris: C.I.D., pp. 894-907.

Vakkari, P. (2001) A theory of the task-based information retrieval process: A summary and generalisation of a longitudinal study. *Journal of Documentation*, 57 (1), pp. 44-60.

Vakkari, P. (2002) Subject knowledge, source of terms, and term selection in query expansion: An analytical study, *Advances in Information Retrieval. Proceedings of 24th BCS-IRSG European Colloquium on IR Research.* Lecture Notes in Computer Science Vol.2291, 2002. Berlin: Springer, pp. 110-123.

Vakkari, P., Pennanen, M. and Serola, S. (2003) Changes of search terms and tactics while writing a research proposal: A longitudinal case study. *Information Processing and Management*, 39 (3), pp. *445-463.*

Voorhees, E. M. (1994) Query expansion using lexical-semantic relations. Croft, W. B. and Rijsbergen, C. J. (eds.) *Proceedings of the Association for Computing Machinery Special Interest Group on Information Retrieval (ACM/SIGIR), 17th Annual International Conference on Research and Development in Information Retrieval*. Berlin: Springer-Verlag, pp. 61-69.

Wilson, T.D. (1981) On user studies and information needs. *Journal of Documentation*, 37 (1), pp. 3-15.

Wilson, T.D. (1999) Models in information behaviour research. *Journal of Documentation*, 55 (3), pp. 249-270.

Appendix A: Invitation Letter

Dear veterinary academic/researcher,

Do you use CAB Abstracts or some other bibliographic database to gain access to current research articles? Are you interested in learning a bit more about using such systems more effectively? Do you have some specific research topics you would like to query against such a system but can't find the time? If the answer to one or more of the above is "yes" then it may be worth your while taking a minute to consider the request to get involved in a research project currently taking place within the Vet School.

As you are all very busy people there is little point in approaching scientists for their time without offering something in return. So, if you are prepared to spend 40-50 minutes getting involved in the short research experiment outlined below I can offer in return an up-to-date bibliographic record set for three topics of your choosing as well as some 'tutorial' input to help you make more effective use of the facilities offered with the CAB Abstracts database. And so to the detail…

I am a doctoral student at the Department of Computer and Information Sciences at the University of Strathclyde working on my dissertation under the supervision of Crawford Revie. As part of my research, I intend to examine how researchers in a range of life science disciplines interact with the thesaurus (subject terminology) of CAB Abstracts, one of the world largest agricultural databases. Specifically, I am trying to understand how researchers' knowledge of a subject domain affects their selection of terms from a thesaurus while searching databases. My research also explores the extent to which the thesaurus aids the selection of alternative search terms.

I carried out a pilot experiment to test a research methodology in which a number of staff from the Faculty of Veterinary Medicine took place. Based on their positive feedback and having made some suggested changes in the format I plan to carry out the full study during November and December of this year. Your involvement would consist of one face-to-face meeting in the Vet School taking no more than one hour.

Should you agree to participate you would be asked to:

1. Complete a pre-search questionnaire which asked for some general background information as well as assessing your level of previous exposure to bibliographic database use (this should take around 10 minutes);

2. During the face-to-face meeting use a Web-based system made available at the time to search for up to three topics of interest to you. This would involve:
   - Searching the CAB Abstracts database for the topic(s)
   - evaluating the records retrieved as a result of your search(s)
   - a brief interview on overall impressions after completing the exercise

Assuming that the topics for which you search result in useful result sets I would also ensure that you received full bibliographic details in a suitable format (paper or electronic).

Any information obtained will of course be kept strictly confidential and will be used for no purpose other than the present investigation. I trust that you will be interested in participating in this project. If so, could you reply to this email by Wednesday 31st October and I will send a list of time slots available for your participation by return.

Appendix B: Sample of Participant Summary Sheet

Please check (or if currently unspecified, provide) your contact details:

Name:

Address:

Phone:

E-mail:

I plan to carry out my experiments during late November and December, and am currently filling out my schedule for the first three weeks of this period. If there is a specific day/time that would be particularly suitable between November 19th – November 30th please indicate this below and I will attempt to schedule you in that slot.

Date: Time:

Just in case it is not possible to allocate you this precise slot could you please indicate below 2-3 alternative times at which you might be available to participate:

	Week commencing 19th		**Week commencing 26th**	
	AM	PM	AM	PM
Monday				
Tuesday				
Wednesday				
Thursday				
Friday				

Contact address:
Ali Asghar Shiri
Department of Computer and Information Sciences
University of Strathclyde
26 Richmond Street
Livingstone Tower
Glasgow G1 1XH UK
Phone: 141- 548 4065 or 141-548 3700
E-mail: shiri@dis.strath.ac.uk

Appendix C: A Tutorial on the Ovid CAB Database

The purpose of this short tutorial is to provide you with a brief introduction to the Ovid CAB database and the steps I would like you to take during a practice search session.

Ovid is one of the leading bibliographic database providers, which makes the CAB Abstracts database, the largest agricultural bibliographic database, available on the Web. Ovid provides alternative ways of searching the database. One of the advanced and unique search facility of the Ovid CAB is an option called "Map term to subject heading" on the search page through which the user' s search term is matched to a thesaurus (domain-specific terminology) of the database for more precise search results. The user will have the opportunity to select terms from the thesaurus. The thesaurus is a structured terminology, which provides the user with alternative search terms such as terms broader, narrower, or synonym to the user's term.

I intend to evaluate the ways in which users interact with the thesaurus and select search terms for carrying out their searches.

Therefore, I would like you to take the following steps to perform the searches:

1. Tick the box besides "Map term to subject heading" on the main search page for initiating you search

2. Enter your search term into the search bar. Note that in order to get better results you are advised to enter your search terms one at a time and combine them at the end of the search. For example if you search

is about " contamination of beef carcass in abattoir ", you will need to search "contamination", "beef", "carcass" and "abattoir" separately and combine them all at the end.

3. Following pressing the search button you will be provided with a number of terms from the thesaurus which you can select as appropriate to your search. You can also click on the hypertext terms to view the details of that specific term such as its broader, narrower or synonym terms.

4. Once you obtain the results you may view and /or reformulate your search depends on how much you are satisfied with the number and content of the records retrieved.

5. At the end of the search you will be provided with printed copies of the results of your search to evaluate their relevance to your search request.

I will record the entire search (i.e. terms entered, thesaurus terms selected, results viewed etc.) so that I can later analyse the whole search process for evaluation of the thesaurus-based browsing and searching and its impact on search term selection.

Appendix D: Pre-search Questionnaire

Personal information

1.1. What is your gender?

- ○ Male
- ○ Female

1.2. Into which age-band do you fall?

- ☐ Less than 30
- ☐ 30 - 40
- ☐ 40 - 50
- ☐ More than 50

1.3. Which category best describes your status?

- ☐ Postgraduate/Doctoral student
- ☐ Faculty/Researcher
- ☐ Other (Please specify)

1.4. What is your educational background?

(subject of degree(s))

BSc: _____

MSc: _____

MPhil: _____

PhD: _____

Computer and search experience

2.1. How long have you been making regular use of computers?

- ☐ Less than 1 year
- ☐ 1 - 3 years
- ☐ 4 - 10 years
- ☐ More than 10 years

2.2. Which of the following computer-related applications do you regularly use?

- ☐ Word processors
- ☐ World Wide Web
- ☐ Other specific applications (such as E-mail, spreadsheets etc.)
- ☐ Electronic version of library catalogues
- ☐ Scientific and technical databases (e.g. Dialog, Datastar, SilverPlatter, etc.)

2.3. Are you familiar with agricultural databases such as CAB Abstracts?

- ○ Yes
- ○ No

2.4. How often do you use CAB Abstracts?

- ☐ I have never used CAB Abstracts
- ☐ 1-2 times a month
- ☐ More than 3 times a month

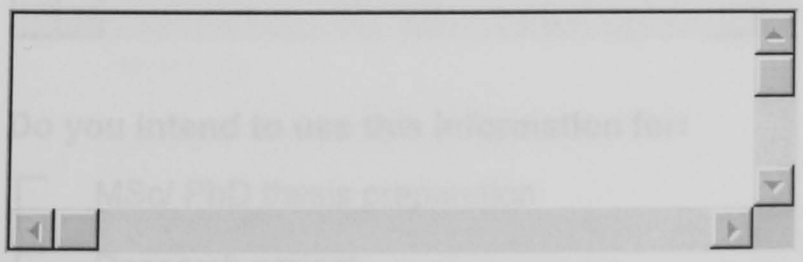
2.5. If you use CAB Abstracts, which of the following versions do you mostly use?

- ☐ CABDirect on the web
- ☐ Dialog CAB Abstracts on the web
- ☐ OVID CAB Abstracts on the web
- ☐ Silver Platter CAB Abstracts on CD-ROM
- ☐ Silver Platter CAB Abstracts on the web

Search characteristics and intention (First Topic)

3.1. During my experiment I would like to observe your searching behaviour using CAB Abstarcts. On the subsequent pages please specify up to three topics of interest to you(1 per page)

First topic



3.2. Do you intend to use this information for:

- ☐ MSc/ PhD thesis preparation
- ☐ Research project
- ☐ Teaching
- ☐ Academic journal preparation
- ☐ Other (please specify)

3.3. How familiar are you with this topic?

☐ Unfamiliar ☐ Moderately familiar ☐ Very familiar

3.4. What type of search will you carry out on this topic?

- ○ Broad search
- ○ Specific search

3.5. Have you previously performed searches on this topic?

- ○ Yes
- ○ No

3.6. If yes, how long ago did you make such a search?

- ☐ 1-6 months ago
- ☐ 7-12 months ago
- ☐ More than one year ago

Second topic

4.1. Please describe the second topic you would like to search for in the database.

Second topic

4.2. Do you intend to use this information for:

- ☐ MSc/ PhD thesis preparation
- ☐ Research project

- ☐ Teaching
- ☐ Academic journal preparation
- ☐ Other (please specify)

4.3. How familiar are you with this topic?

☐ Unfamiliar ☐ Moderately familiar ☐ Very familiar

4.4. What type of search will you carry out on this topic?

- ○ Broad search
- ○ Specific search

4.5. Have you previously performed searches on this topic?

- ○ Yes
- ○ No

4.6. If yes, how long ago did you make such a search?

- ☐ 1-6 months ago
- ☐ 7-12 months ago
- ☐ More than one year ago

Third topic

5.1. Please describe the third topic you would like to search for in the database.

Third topic

5.2. Do you intend to use this information for:

- ☐ MSc/ PhD thesis preparation
- ☐ Research project
- ☐ Teaching
- ☐ Academic journal preparation
- ☐ Other(please specify)

5.3. How familiar are you with this topic?

☐ Unfamiliar ☐ Moderately familiar ☐ Very familiar

5.4. What type of search would like to do?

○ Broad search

○ Specific search

5.5. Have you previously performed searches on this topic?

○ Yes

○ No

5.6. If yes, how long ago did you make such a search?

☐ 1-6 months ago

☐ 7- 12 months ago

☐ More than one year ago

Appendix E: Post-search Questionnaire

Search Term Satisfaction

1.1. Did the thesaurus provide you with additional search terms for selection?

- ○ Yes
- ○ No

1.2. If yes, did you find the alternative terms you selected during the thesaurus browsing useful?

- ○ Yes
- ○ No

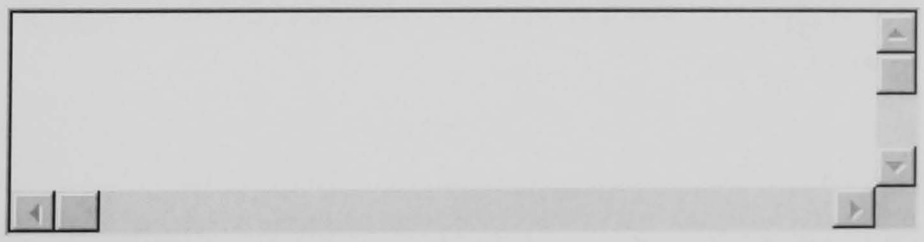
1.3. Were there additional terms suggested that you were not aware of at the beginning of the search?

- ○ Yes
- ○ No

1.4. How close were the terms you selected from the thesaurus to your original search topic?

○ Very close ○ Fairly close ○ Not at all

1.5. What caused you to select the additional terms?



1.6. How satisfied were you with the number of terms you selected from the thesaurus?

- ○ Very satisfied
- ○ Fairly satisfied
- ○ Unsatisfied
- ○ Very unsatisfied

Search Results Satisfaction

2.1. How satisfied are you with the number of references in the search results?

- ○ Very satisfied
- ○ Fairly satisfied
- ○ Unsatisfied
- ○ Very unsatisfied

2.2. Based on the bibliographic information, how relevant did the documents you retrieved appear to be?

- ○ Very relevant
- ○ Partially relevant
- ○ Not relevant

2.3. Do you think the results you viewed generally matched the topic you were interested in?

- ○ Very good match
- ○ Fairly good match
- ○ Poor match
- ○ Not at all

Appendix F: Post-session Interview Script

Thesaurus

1. **How easy was the thesaurus browsing and navigation?**
   (Where 1 is very difficult and 5 is very easy)

   ○ ○ ○ ○ ○
   1 2 3 4 5

2. **What, if any, effect did browsing the thesaurus have on your selection of search terms?**

3. **What kinds of problems or difficulties did you experience during the thesaurus browsing?**

4. **What type of help do you think you needed during browsing the thesaurus?**

Interface

5. **How east was to learn and use the interface?**
   (Where 1 is very difficult and 5 is very easy)

   ○ ○ ○ ○ ○
   1 2 3 4 5

6. **How easy was to use commands on the interface such as "continue" etc.?**
   (Where 1 is very difficult and 5 is very easy)

   ○ ○ ○ ○ ○
   1 2 3 4 5

7. **What did you like or dislike about the interface?**

Thesaurus Experience

8. **Before searching this database did you know that the CAB abstracts database a thesaurus?**

○ Yes ○ No

9. **Have you ever used the CAB thesaurus?**

○ Yes ○ No

10. **Have you ever used any thesaurus for searching?**

○ Yes ○ No

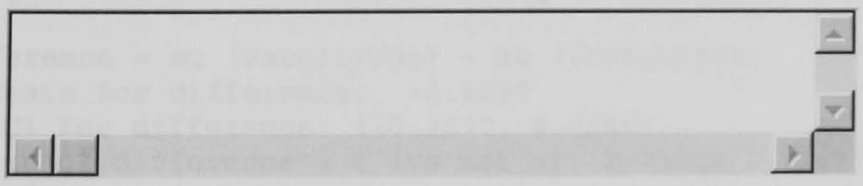
11. **If you have used CAB or other thesauri, how familiar were you with the thesaurus?**

(Where 1 is very difficult and 5 is very easy)

○ ○ ○ ○ ○

1 2 3 4 5

12. **Do you have any other comments about the thesaurus, interface or the search in general?**



Appendix G: Results of Statistical Tests

The following statistical test results are arranged based on the order of hypotheses whose results reported in Chapter 5.

Two-Sample T-Test and CI: Cognitive moves, Status (Table 5.1)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   13     29.08      5.24       1.5
Postgradua   15     29.47      6.23       1.6

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  -0.39
95% CI for difference: (-4.85, 4.07)
T-Test of difference = 0 (vs not =): T-Value = -0.18  P-Value = 0.859
DF = 25
```

Two-Sample T-Test and CI: Selection, Status (Table 5.2)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   42     3.02      1.33       0.21
Postgradua   44     2.55      1.04       0.16

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  0.478
95% CI for difference: (-0.034, 0.991)
T-Test of difference = 0 (vs not =): T-Value = 1.86  P-Value = 0.067
DF = 84
Both use Pooled StDev = 1.19
```

Two-Sample T-Test and CI: Browsing Titles, Status (Table 5.2)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   42     1.143     0.354      0.055
Postgradua   44     1.273     0.544      0.082

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  -0.1299
95% CI for difference: (-0.3277, 0.0680)
T-Test of difference = 0 (vs not =): T-Value = -1.31  P-Value = 0.195
DF = 84
Both use Pooled StDev = 0.461
```

Two-Sample T-Test and CI: Reformulation, Status (Table 5.2)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   42     1.071     0.745      0.12
Postgradua   44     1.114     0.784      0.12

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  -0.042
95% CI for difference: (-0.371, 0.286)
T-Test of difference = 0 (vs not =): T-Value = -0.26  P-Value = 0.799
DF = 84
Both use Pooled StDev = 0.765
```

Two-Sample T-Test and CI: Operators, Status (Table 5.2)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   42     1.833     0.986      0.15
Postgraduate  44      2.07      1.13      0.17

Difference = mu (Faculty/Re) - mu (Postgraduate)
Estimate for difference:  -0.235
95% CI for difference: (-0.690, 0.220)
T-Test of difference = 0 (vs not =): T-Value = -1.03  P-Value = 0.308
DF = 84
Both use Pooled StDev = 1.06
```

Two-Sample T-Test and CI: Initial Terms, Status (Table 5.2)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   42      2.81      1.33      0.21
Postgradua   44      2.48      1.07      0.16

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  0.332
95% CI for difference: (-0.184, 0.848)
T-Test of difference = 0 (vs not =): T-Value = 1.28  P-Value = 0.204
DF = 84
Both use Pooled StDev = 1.20
```

Two-Sample T-Test and CI: Physical, Status (Table 5.3)

```
Two-sample T for Physical

Status        N      Mean     StDev   SE Mean
Faculty/Re   13      58.3      14.1       3.9
Postgradua   14      58.0      14.4       3.8

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  0.31
95% CI for difference: (-11.01, 11.63)
T-Test of difference = 0 (vs not =): T-Value = 0.06  P-Value = 0.956  DF = 24
```

Two-Sample T-Test and CI: Perform Search, Status (Table 5.4)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   44      3.00      1.45      0.22
Postgradua   44      2.55      1.25      0.19

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  0.455
95% CI for difference: (-0.118, 1.027)
T-Test of difference = 0 (vs not =): T-Value = 1.58  P-Value = 0.118
DF = 84
```

Two-Sample T-Test and CI: Scroll Up & Down, Status (Table 5.4)

```
Two-sample T for Scroll Up & down

Status        N      Mean     StDev   SE Mean
Faculty/Re   44      7.75      5.29      0.80
Postgradua   44      6.45      5.57      0.84

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  1.30
95% CI for difference: (-1.01, 3.60)
T-Test of difference = 0 (vs not =): T-Value = 1.12  P-Value = 0.266
DF = 85
```

Two-Sample T-Test and CI: Back & Forward, Status (Table 5.4)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   44     0.227     0.803      0.12
Postgradua   44      0.61      1.30      0.20

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  -0.386
95% CI for difference: (-0.845, 0.072)
T-Test of difference = 0 (vs not =): T-Value = -1.68  P-Value = 0.098
DF = 71
```

Two-Sample T-Test and CI: Continue, Status (Table 5.4)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   44      3.82      1.91      0.29
Postgradua   44      3.36      1.93      0.29

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  0.455
95% CI for difference: (-0.359, 1.268)
T-Test of difference = 0 (vs not =): T-Value = 1.11  P-Value = 0.270
DF = 85
```

Two-Sample T-Test and CI: Combine, Status (Table 5.4)

Status	N	Mean	StDev	SE Mean
Faculty/Re	44	1.95	1.12	0.17
Postgradua	44	2.09	1.12	0.17

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference: -0.136
95% CI for difference: (-0.610, 0.338)
T-Test of difference = 0 (vs not =): T-Value = -0.57 P-Value = 0.569
DF = 85

Two-Sample T-Test and CI: Display & Email, Status (Table 5.4)

Status	N	Mean	StDev	SE Mean
Faculty/Re	44	2.932	0.846	0.13
Postgradua	44	3.36	1.26	0.19

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference: -0.432
95% CI for difference: (-0.887, 0.024)
T-Test of difference = 0 (vs not =): T-Value = -1.89 P-Value = 0.063
DF = 75

Two-Sample T-Test and CI: Main Search Page, Status (Table 5.4)

Status	N	Mean	StDev	SE Mean
Faculty/Re	44	0.114	0.321	0.048
Postgradua	44	0.091	0.291	0.044

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference: 0.0227
95% CI for difference: (-0.1071, 0.1526)
T-Test of difference = 0 (vs not =): T-Value = 0.35 P-Value = 0.729
DF = 85

Two-Sample T-Test and CI: Next & Previous Page, Status (Table 5.4)

Status	N	Mean	StDev	SE Mean
Faculty/Re	44	0.77	1.01	0.15
Postgradua	44	0.70	1.00	0.15

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference: 0.068
95% CI for difference: (-0.358, 0.494)
T-Test of difference = 0 (vs not =): T-Value = 0.32 P-Value = 0.751
DF = 85

Two-Sample T-Test and CI: Search History, Status (Table 5.4)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   44     0.182     0.390     0.059
Postgradua   44     0.114     0.321     0.048

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  0.0682
95% CI for difference: (-0.0833, 0.2197)
T-Test of difference = 0 (vs not =): T-Value = 0.90  P-Value = 0.373
DF = 82
```

Two-Sample T-Test and CI: Contract & Expand, Status (Table 5.4)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   44     0.386     0.655     0.099
Postgradua   44     0.341     0.776      0.12

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  0.045
95% CI for difference: (-0.259, 0.350)
T-Test of difference = 0 (vs not =): T-Value = 0.30  P-Value = 0.767
DF = 83
```

Two-Sample T-Test and CI: Terms Browsed, Status (Table 5.5)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   15     130.8      55.7        14
Postgradua   15      96.2      37.3       9.6

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  34.6
95% CI for difference: (-1.1, 70.3)
T-Test of difference = 0 (vs not =): T-Value = 2.00  P-Value = 0.057
DF = 24
```

Two-Sample T-Test and CI: Terms Selected, Status (Table 5.6)

```
Status        N      Mean     StDev   SE Mean
Faculty/Re   15     16.93      7.51       1.9
Postgradua   15     15.27      6.08       1.6

Difference = mu (Faculty/Re) - mu (Postgradua)
Estimate for difference:  1.67
95% CI for difference: (-3.46, 6.79)
T-Test of difference = 0 (vs not =): T-Value = 0.67  P-Value = 0.510
DF = 26
```

Two-Sample T-Test and CI: Cognitive moves, Gender (Table 5.7)

Gender	N	Mean	StDev	SE Mean
Female	16	27.44	4.84	1.2
Male	12	31.75	5.99	1.7

```
Difference = mu (Female) - mu (Male  )
Estimate for difference:  -4.31
95% CI for difference: (-8.71, 0.09)
T-Test of difference = 0 (vs not =): T-Value = -2.04  P-Value = 0.054
DF = 20
```

Two-Sample T-Test and CI: Physical Moves, Gender (Table 5.7)

Gender	N	Mean	StDev	SE Mean
Female	16	54.6	10.9	2.7
Male	11	63.4	16.7	5.0

```
Difference = mu (Female) - mu (Male  )
Estimate for difference:  -8.80
95% CI for difference: (-20.99, 3.39)
T-Test of difference = 0 (vs not =): T-Value = -1.54  P-Value = 0.145
DF = 15
```

Two-Sample T-Test and CI: Physical, Experience Level (Table 5.8)

EXP_Leve	N	Mean	StDev	SE Mean
Experience	19	57.9	12.9	3.0
Novice	8	58.8	17.2	6.1

```
Difference = mu (Experience) - mu (Novice    )
Estimate for difference:  -0.86
95% CI for difference: (-15.92, 14.21)
T-Test of difference = 0 (vs not =): T-Value = -0.13  P-Value = 0.902
DF = 10
```

Two-Sample T-Test and CI: Cognitive, Experience Level (Table 5.8)

Expe_Lev	N	Mean	StDev	SE Mean
Experience	21	29.52	5.75	1.3
Novice	7	28.57	5.88	2.2

```
Difference = mu (Experience) - mu (Novice)
Estimate for difference:  0.95
95% CI for difference: (-4.74, 6.64)
T-Test of difference = 0 (vs not =): T-Value = 0.37  P-Value = 0.717
DF = 10
```

Two-Sample T-Test and CI: Cognitive, Complexity (Table 5. 9)

```
Two-sample T for Cognitive

Complexity   N      Mean      StDev   SE Mean
1           51      9.88       3.30      0.46
2           35     16.23       3.69      0.62

Difference = mu (1) - mu (2)
Estimate for difference:  -6.346
95% CI for difference: (-7.858, -4.835)
T-Test of difference = 0 (vs not =): T-Value = -8.35  P-Value = 0.000
DF = 84
Both use Pooled StDev = 3.46
```

Two-Sample T-Test and CI: Physical, Topic Complexity (Table 5.10)

```
Two-sample T for Physical

TopicCom     N      Mean      StDev   SE Mean
Complex     37     24.70       6.65       1.1
Simple      51     17.29       7.65       1.1

Difference = mu (Complex) - mu (Simple )
Estimate for difference:  7.41
95% CI for difference: (4.36, 10.45)
T-Test of difference = 0 (vs not =): T-Value = 4.84  P-Value = 0.000
DF = 83
```

Two-Sample T-Test and CI: Reformulation, Topic Complexity (Table 5.11)

```
TopicCom     N      Mean      StDev   SE Mean
Complex     39     1.513      0.854      0.14
Simple      51     0.882      0.653     0.091

Difference = mu (Complex) - mu (Simple )
Estimate for difference:  0.630
95% CI for difference: (0.302, 0.959)
T-Test of difference = 0 (vs not =): T-Value = 3.83  P-Value = 0.000  DF = 69
```

Two-Sample T-Test and CI: Terms Selected, Topic Complexity (Table 5.12)

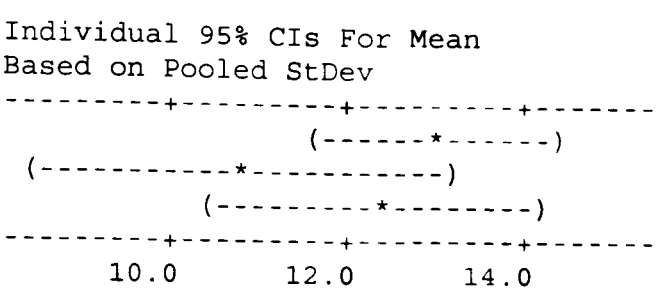
```
TopicCom     N      Mean      StDev   SE Mean
Complex     39      7.74       3.51      0.56
Simple      51      3.55       1.70      0.24

Difference = mu (Complex) - mu (Simple )
Estimate for difference:  4.195
95% CI for difference: (2.970, 5.419)
T-Test of difference = 0 (vs not =): T-Value = 6.88  P-Value = 0.000  DF = 51
```

One-way ANOVA: Cognitive, Topic Familiarity (Table 5.13)

Source	DF	SS	MS	F	P
TopicFam	2	53.8	26.9	1.25	0.292
Error	83	1789.6	21.6		
Total	85	1843.4			

Level	N	Mean	StDev
Moderate	49	13.000	4.878
Unfamili	14	10.786	3.577
Very fam	23	12.348	4.677

Pooled StDev = 4.643

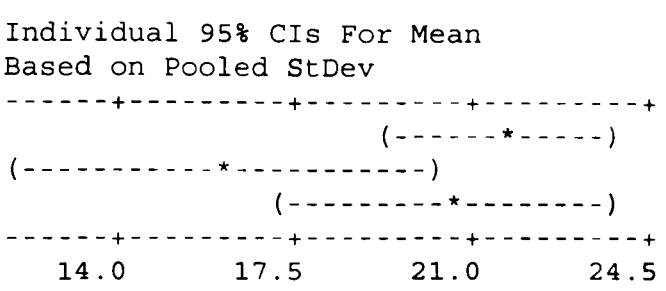


One-way ANOVA: Physical Moves, Topic Familiarity (Table 5.13)

Source	DF	SS	MS	F	P
TopicFam	2	350.3	175.1	2.78	0.067
Error	85	5347.0	62.9		
Total	87	5697.3			

Level	N	Mean	StDev
Moderate	50	21.640	8.827
Unfamili	15	16.133	6.151
Very fam	23	20.522	6.741

Pooled StDev = 7.931



Two-Sample T-Test and CI: Cognitive Moves, Topic Experience (Table 5.14)

TopicExp	N	Mean	StDev	SE Mean
FALSE	31	12.10	5.49	0.99
TRUE	55	12.67	4.16	0.56

Difference = mu (FALSE) - mu (TRUE)
Estimate for difference: -0.58
95% CI for difference: (-2.66, 1.51)
T-Test of difference = 0 (vs not =): T-Value = -0.55 P-Value = 0.585
DF = 84
Both use Pooled StDev = 4.68

Two-Sample T-Test and CI: Physical Moves, Topic Experience (Table 5.14)

TopicExp	N	Mean	StDev	SE Mean
FALSE	33	19.82	7.50	1.3
TRUE	55	20.76	8.47	1.1

Difference = mu (FALSE) - mu (TRUE)
Estimate for difference: -0.95
95% CI for difference: (-4.40, 2.51)
T-Test of difference = 0 (vs not =): T-Value = -0.54 P-Value = 0.587
DF = 74

Two-Sample T-Test and CI: Physical Moves, Search Type (Table 5.15)

```
SearchTy      N      Mean     StDev   SE Mean
Broad sear   55     21.29      8.53       1.2
Specific s   33     18.94      7.19       1.3

Difference = mu (Broad sear) - mu (Specific s)
Estimate for difference:  2.35
95% CI for difference: (-1.03, 5.74)
T-Test of difference = 0 (vs not =): T-Value = 1.38  P-Value = 0.171
DF = 76
```

Two-Sample T-Test and CI: Cognitive Moves, Search Type (Table 5.15)

```
SearchTy      N      Mean     StDev   SE Mean
Broad sear   54     11.94      4.61      0.63
Specific s   32     13.34      4.67      0.83

Difference = mu (Broad sear) - mu (Specific s)
Estimate for difference:  -1.40
95% CI for difference: (-3.46, 0.66)
T-Test of difference = 0 (vs not =): T-Value = -1.35  P-Value = 0.180
DF = 84
Both use Pooled StDev = 4.63
```

One-way ANOVA: Terms Browsed, Topic Familiarity (Table 5.16)

```
Source      DF       SS       MS        F       P
TopicFam     2     3748     1874     3.75   0.027
Error       87    43483      500
Total       89    47231

                                                 Individual 95% CIs For Mean
                                                 Based on Pooled StDev
Level       N     Mean    StDev  -+---------+---------+---------+-----
Moderate   52    40.27    23.19                      (-----*----)
Unfamili   15    23.47    13.65           (---------*--------)
Very fam   23    41.70    24.73                     (-------*------)
                                 -+---------+---------+---------+-----
Pooled StDev =    22.36          12        24        36        48
```

One-way ANOVA: Terms Selected, Topic Familiarity (Table 5.16)

```
Source      DF       SS       MS        F       P
TopicFam     2     16.4      8.2     0.73   0.486
Error       87    984.5     11.3
Total       89   1000.9

                                                 Individual 95% CIs For Mean
                                                 Based on Pooled StDev
Level       N     Mean    StDev  ---------+---------+---------+------
Moderate   52    5.731    3.593                      (--------*---------)
Unfamili   15    4.800    2.731  (-----------------*-----------------)
Very fam   23    4.913    3.175      (-------------*-------------)
                                 ---------+---------+---------+------
Pooled StDev =    3.364                 4.0       5.0       6.0
```

Chi-Square Test: Topic Experience, Additional Terms (Table 5.17)

	No exper	Experien	Total
1	1	15	16
	5.87	10.13	
2	32	42	74
	27.13	46.87	
Total	33	57	90

Chi-Sq = 4.037 + 2.337 + 0.873 + 0.505 = 7.753
DF = 1, P-Value = 0.005

Chi-Square Test: Term Usefulness and Topic Experience (Table 5.18)

	term use	term not	Total
1	4	29	33
	9.17	23.83	
2	21	36	57
	15.83	41.17	
Total	25	65	90

Chi-Sq = 2.912 + 1.120 + 1.686 + 0.648 = 6.367
DF = 1, P-Value = 0.012

Two-Sample T-Test and CI: Cognitive Moves, Result Relevance (Table 5.19)

ResultRe	N	Mean	StDev	SE Mean
Partially	27	13.37	4.71	0.91
Very relev	59	12.05	4.61	0.60

Difference = mu (Partially) - mu (Very relev)
Estimate for difference: 1.32
95% CI for difference: (-0.83, 3.47)
T-Test of difference = 0 (vs not =): T-Value = 1.22 P-Value = 0.225
DF = 84

Two-Sample T-Test and CI: Physical Moves, Result Relevance (Table 5.19)

ResultRe	N	Mean	StDev	SE Mean
Partially	29	23.45	8.04	1.5
Very relev	59	18.92	7.75	1.0

Difference = mu (Partially) - mu (Very relev)
Estimate for difference: 4.53
95% CI for difference: (0.92, 8.15)
T-Test of difference = 0 (vs not =): T-Value = 2.51 P-Value = 0.015
DF = 53

Two-Sample T-Test and CI: Cognitive Moves, Term Usefulness (Table 5.20)

```
TermUsef      N      Mean      StDev    SE Mean
FALSE        24     10.83       3.89       0.79
TRUE         62     13.10       4.80       0.61

Difference = mu (FALSE) - mu (TRUE )
Estimate for difference:  -2.26
95% CI for difference: (-4.45, -0.08)
T-Test of difference = 0 (vs not =): T-Value = -2.06  P-Value = 0.043
 DF = 84
Both use Pooled StDev = 4.57
```

Two-Sample T-Test and CI: Physical Moves, Term Usefulness (Table 5.20)

```
TermUsef      N      Mean      StDev    SE Mean
FALSE        25     18.92       8.09        1.6
TRUE         63     21.00       8.08        1.0

Difference = mu (FALSE) - mu (TRUE )
Estimate for difference:  -2.08
95% CI for difference: (-5.93, 1.77)
T-Test of difference = 0 (vs not =): T-Value = -1.09  P-Value = 0.283
DF = 44
```

Two-Sample T-Test and CI: Terms Selected, Term Usefulness (Table 5.21)

```
TermUsef      N      Mean      StDev    SE Mean
FALSE        25     4.08       2.22       0.44
TRUE         65     5.86       3.59       0.45

Difference = mu (FALSE) - mu (TRUE )
Estimate for difference:  -1.782
95% CI for difference: (-3.035, -0.528)
T-Test of difference = 0 (vs not =): T-Value = -2.83  P-Value = 0.006
DF = 70
```

Two-Sample T-Test and CI: Terms Selected, Result Relevance (Table 5.22)

```
ResultRe      N      Mean      StDev    SE Mean
Partially    30     6.40       3.85       0.70
Very relev   60     4.85       2.98       0.38

Difference = mu (Partially ) - mu (Very relev)
Estimate for difference:  1.550
95% CI for difference: (-0.062, 3.162)
T-Test of difference = 0 (vs not =): T-Value = 1.94  P-Value = 0.059  DF = 46
```

Two-Sample T-Test and CI: Cognitive, Time Per Concept (Table 5.23)

```
ConcepTi      N      Mean     StDev   SE Mean
Less than    19     27.47      4.67       1.1
More than     9     33.11      5.99       2.0

Difference = mu (Less than ) - mu (More than )
Estimate for difference:  -5.64
95% CI for difference: (-10.57, -0.70)
T-Test of difference = 0 (vs not =): T-Value = -2.49  P-Value = 0.029  DF = 12
```

Chi-Square Test: Ease of Learning/Use and Ease of Browsing (Table 5.24)

```
           Easy Very eas    Total
    1        16        3       19
          12.67     6.33

    2         4        7       11
           7.33     3.67

Total        20       10       30

Chi-Sq =  0.877 +  1.754 + 1.515 +  3.030 = 7.177
DF = 1, P-Value = 0.007
1 cells with expected counts less than 5.0
```

Two-Sample T-Test and CI: Physical Moves, Ease of Browsing/Navigation (Table 5.25)

```
EaseOfBr      N      Mean     StDev   SE Mean
Easy         16      64.1      14.0       3.5
Very easy    11     49.45      8.65       2.6

Difference = mu (Easy      ) - mu (Very easy )
Estimate for difference:  14.67
95% CI for difference: (5.67, 23.67)
T-Test of difference = 0 (vs not =): T-Value = 3.37  P-Value = 0.003
DF = 24
```

Two-Sample T-Test and CI: Cognitive Moves, Ease of Browsing/Navigation (Table 5.26)

```
EaseOfBr      N      Mean     StDev   SE Mean
Easy         17     30.24      6.19       1.5
Very easy    11     27.82      4.71       1.4

Difference = mu (Easy      ) - mu (Very easy )
Estimate for difference:  2.42
95% CI for difference: (-1.84, 6.67)
T-Test of difference = 0 (vs not =): T-Value = 1.17  P-Value = 0.253
DF = 25
```

Two-Sample T-Test and CI: Terms Selected, Ease of Browsing/Navigation
(Table 5.27)

```
EaseBrow      N      Mean     StDev   SE Mean
2            19     16.47      7.36       1.7
3            11     15.45      5.87       1.8

Difference = mu (2) - mu (3)
Estimate for difference:  1.02
95% CI for difference: (-4.03, 6.07)
T-Test of difference = 0 (vs not =): T-Value = 0.42  P-Value = 0.681
DF = 24
```

Two-Sample T-Test and CI: Terms Browsed, Ease of Browsing/Navigation
(Table 5.27)

```
EaseBrow      N      Mean     StDev   SE Mean
2            18     119.2      50.4        12
3            11      93.5      31.9       9.6

Difference = mu (2) - mu (3)
Estimate for difference:  25.7
95% CI for difference: (-5.7, 57.1)
T-Test of difference = 0 (vs not =): T-Value = 1.68  P-Value = 0.105
DF = 26
```

Two-Sample T-Test and CI: Reformulation, Ease of Browsing/Navigation
(Table 5.28)

```
EaseBrow      N      Mean     StDev   SE Mean
2            19      1.58      1.54      0.35
3            11     0.545     0.522      0.16

Difference = mu (2) - mu (3)
Estimate for difference:  1.033
95% CI for difference: (0.236, 1.831)
T-Test of difference = 0 (vs not =): T-Value = 2.67  P-Value = 0.013  DF = 24
```

Two-Sample T-Test and CI: Time Per Search, Ease of Browsing/Navigation
(Table 5.29)

```
EaseBrow      N      Mean     StDev   SE Mean
2            19     23.37      8.67       2.0
3            11     16.82      5.34       1.6

Difference = mu (2) - mu (3)
Estimate for difference:  6.55
95% CI for difference: (1.30, 11.80)
T-Test of difference = 0 (vs not =): T-Value = 2.56  P-Value = 0.016  DF = 27
```

Appendix H: List of Search Topics Provided by Users

N	Search Topics
1	Osteochondrosis in horses
2	Endochondral ossification regulation
3	Inter and Intra - observer Reliability of Interpretation of Diagnostic Imaging
4	Cattle and Listeria
5	Eyes and Listeria
6	Serological Tests
7	Phenotypic Traits of Eggshell Quality
8	Candidate Gene Approach/ Marker Associated Selection
9	Non - Invasive Techniques to Monitor Eggshell Quality
10	Orthopaedics and canine and fracture (PhD)
11	Morphometry and pelvis
12	Cross sectional anatomy
13	Septic arthritis in horses
14	Corrective shoeing in horses
15	Intra articular anaesthesia in horses
16	Porcine endogeu retrovirus
17	Xenotransplantation
18	Cyclosporin
19	Jaundile in cattle
20	Sialocele
21	Mastitis in bovine
22	Veterinary ultrasonography
23	3D ultrasonography (human and veterinary medicine)
24	Canine reproduction and development of embryos and ultrasound of pregnant bitches
25	Pain in horses (The user added "Lameness" after viewing some records)
26	Urolithiasis in horses
27	Cushing's disease in horses
28	Chronic obstructive pulmonary disease Airway obstruction of horses (PhD researcher)
29	Breath condensates Hydro proxide (Measurement was added later in the search process)
30	Asthma risk scanning questionnaires
31	Cyathostomiasis
32	Grass sickness
33	Equine brain tumours
34	Inflammatory mediators in bovine mastitis

N	Search Topics
35	Pharmacodynamics of non-steriodal anti-inflammatory drugs
36	Pain and hyperalgesia in cattle
37	trasngenic and genetically modified mice
38	Clinical conditions and problems with genetically modified mice
39	Assessment of genetically modified mice
40	Hyperalgesia and allodynia
41	Animal pain behaviour
42	Pain sorting schemes
43	Antimicrobial growth promoters
44	Antimicrobial resistance in enterococci farm animals
45	Erythromycin resistance in campylobactoer
46	Bovine respiratory disease (pneumonia)
47	Bovine placenta
48	Mastitis in bovine
49	Microbiological risk assessment
50	Foot and mouth disease modelling (The user mentioned her intention for this search as " general interest" and indicated the "other" category.
51	risk assessment in the area of antimicrobial resistance
52	Feline immunodeficiency virus (FIV)
53	Viral quantification by PCR
54	Cloning dolly (The user mentioned her intention as " general interest".
55	Oncology and telomerase
56	Aging and dogs and telomo
57	Sarloid virus
58	equine ophthalmology
59	Geriatric equine
60	Equine bladder disease
61	vaccines for nematodes in animals (sheep/cattle) (The user mentioned his intention of this search as " research project assessment".
62	E coli 0157 epidemiology intervention and microbiology (The user mentioned his intention of this search as " research project assessment".
63	Exposure to organophosphates (sheep, insect sprays) (The user mentioned his intention of this search as " research project assessment".
64	Runx/CBFA/AML genes in cancer
65	Mouse development
66	Acrosome proteins in mouse spermatozoa
67	Laminitis
68	Tetrology of fallot in horses
69	Acupunture: use in veterinary medicine
70	Epidemiology of thileria parva
71	Diagnosis of bovine trypanosomosis
72	Decision support in animal disease diagnosis

N	Search Topics
73	Equine infalmmatory airway disease/ chronic obstructive airway disease recurrent airway obstruction
74	Equivalence testing/ non-inferiority testing
75	Pain scoring systems/Grade
76	Canine adenovirus
77	Feline cytokines
78	Feline viral vaccines
79	Cardiovascular physiology
80	Equine neurology
81	Equine cushing's disease
82	Mouse mammary tumour virus
83	PMWS/PDNS
84	Bovine dilated cardiomyopathy
85	Canine hyperadrocoticism
86	Human Hyperadrenocorticism: treatment and prognosis
87	Feline leukaemia virus (the user mentioned his intention as " keeping up to date with literature")
88	Diagnostic imaging of the tympanic bullae in the dog and cat
89	Ultrasound imaging of bone
90	Ultrasound hydronephrosis and hydroureter in dogs and cats

Appendix I: List of Initial Search Terms

Topic	Initial Terms Input by Users
1	Osteochodrosis
	Horse
	genetics
2	endochondral ossification
	regulation
	endochondral
3	Diagnostic imaging
	Reliability
	Interpretation
	Reproducibility
	repeatability
4	Cattle
	Listeria
5	Serological tests
6	Eyes
7	Phenotypic
	Eggshell
8	Gene
	Selection
9	Non-invasive
	eggshell
	resonance
	egg
10	Orthopaedics
	Canine
	Fractures
11	Morphometry
	Pelvic
	Pelvic
12	Cross sectional
	Anatomy
	Dogs
13	Septic arthritis
14	Shoeing
	Horse
	Corrective shoeing
15	Intera-articular anasthaesia
16	Porcine
	Endogenous
	Retrovirus
17	Xenotransplantation
	Xenotransplantation
18	Cyclosporin
	Immunosuppresive agents

Topic	Initial Terms Input by Users
	Transplant
	Transplantation
19	Jaundice
20	Sialocele
21	Mastitis
	staphylococcus aureus
	Somatic cell count
22	Ultrasonography
	veterinary
23	Three dimensional
24	Canine
	Reproduction
25	Pain
	Horses
	Lameness
26	Urolithiasis
27	Cushing's
28	Chronic pulmonary disease
	Equine
29	Breath
	Condensate
	Hydrogen peroxide
	Measurement
30	Questionnaire
	Asthma
	Mail
	Survey
31	Cyathostomiasis
32	Grass Sickness
33	Equine
	Brain
	Tumours
	brain
	Brain tumours
34	Bovine
	Mastitis
	Inflammation
35	Pharmacodynamics
	NASID
36	Pain
	Hyperalgesia
	Cattle
	Hyperalgesia
37	Trangenic
	Mouse
38	Mice
	genetic

Topic	Initial Terms Input by Users
	Pathology
	pathology
39	Assessment
	behavioural
	behavioural
40	Hyperalgesia
	Allodynia
41	Animal
	Pain
	Behaviour
	Dog
42	Score
	Pain scoring
	pain
43	Antimicrobial
	Growth promoter
	Resistance
	Pigs
44	Enterococci
	Farm animals
45	Erythromycin
	Resistance
	Campylobacter
46	Bovine
	Respiratory
	Calf
47	Bovine
	Placenta
	Placentome
48	Bovine
	Mastitis
49	Microbiological
	Risk assessment
50	FMD
	Mathematical modelling
	FMD
51	Risk assessment
	Antimicrobial resistance
52	FIV
	Strains
	Strains
53	PCR
	Quantitative
	Viral
54	Dolly
55	Oncology
	Telomerase

Topic	Initial Terms Input by Users
56	Aging
	Dogs
	Telomere
	Review
57	Sarcoid
58	Equine
	Ophthalmology
59	Geriatric
	Eye
60	Equine
	Bladder
	Tumour
61	Vaccines
	Nematodes
	Sheep
	Model
62	E coli
	Intervention
	Epidemiology
	Ecoli 0157
63	Organophospates
	Sheep dips
64	Core binding factor
	Runt
	CBFA1
65	Mouse
	Development
	Microscopy
66	Mice
	Acrosome
	Sperm
	Protein
67	Laminitis
	Horses
	Treatment
68	tetralogy
	Fallot
	Horses
69	Acupuncture
	Horses
	Pain
70	Theileria
	Epidemiology
71	Diagnosis
	Trypanosomiasis
	Trypanosomosis
	Bovine

Topic	Initial Terms Input by Users
72	Decision support
	Animal disease
	Diagnosis
	Expert systems
	Decision
73	Equine
	Chronic
	Obstructive
	Disease
	Airway
74	Equivalence
	Test
	Testing
	Non-inferiority
75	Pain
	Scoring
76	Canine
77	Feline
	Cat
	Cytokine
78	Feline
	Cat
	Vaccines
	Virus
	vaccines
79	Equine
	Cardiology
	Physiology
80	Neurology
	Forebrain
	Lesions
81	Cushings
82	Mouse
	Mammary
	Tumour
	Virus
	MMTV
	Dogs
	Man
83	PMWS
	PMWS
84	PDNS
	Porcine Dermatitis Syndrome
85	Canine
	Hyperadrenocorticism
86	Human
	Hyperadrenocorticism

Topic	Initial Terms Input by Users
	Cushings
87	Feline
88	Diagnostic imaging
	Tympanic bullae
	Dog
	Canine
	Cat
	Feline
89	Ultrasound
	Bone
	Tympanic bullae
	Physics
	Artefacts
	Physical
90	Ultrasound
	Sonography
	Hydronephrosis
	Hydroureter
	Dogs
	Cats
	Hydroureter

Appendix J: Screenshots of the Ovid Search Interface

The following screenshots show various Ovid search interface features. These screenshots were used as transaction logs for analysis.

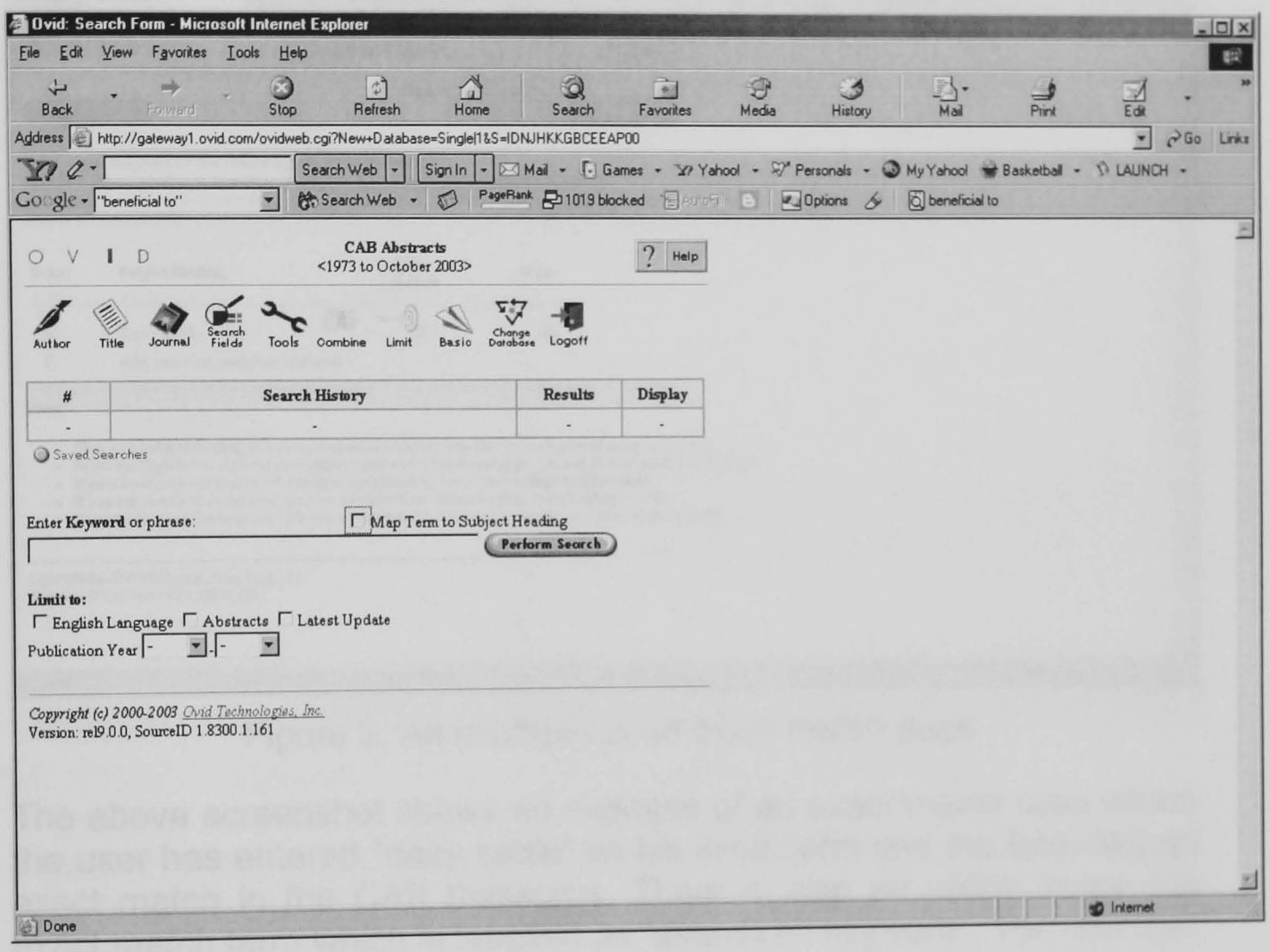


Figure1. Ovid advanced search page

The advanced search page provides users with a facility labelled as “Map Term to Subject Heading” which can be found above the “perform search” button. By ticking the box next to the facility users’ terms are mapped to the terms in the CAB thesaurus.

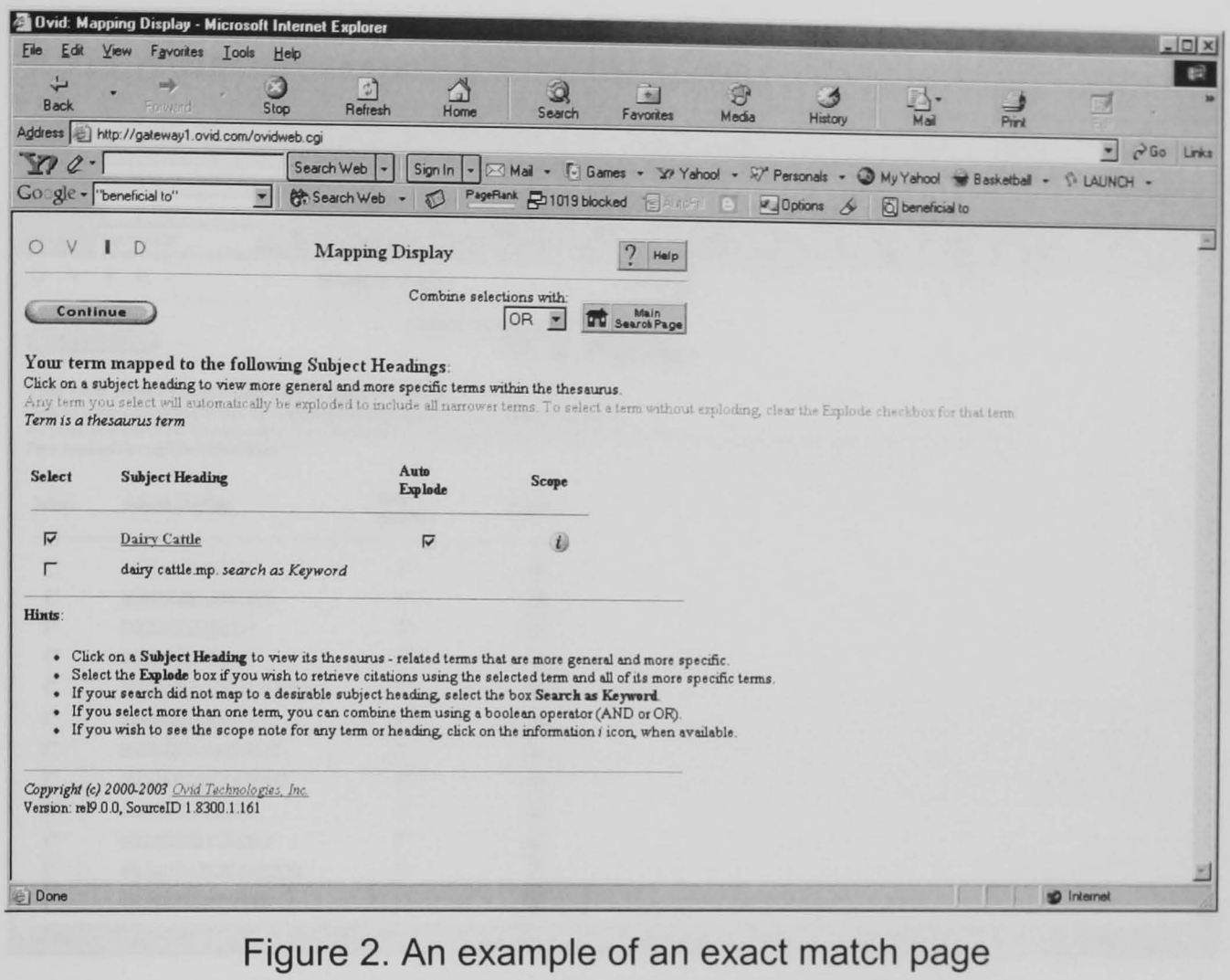


Figure 2. An example of an exact match page

The above screenshot shows an example of an exact match case where the user has entered “dairy cattle” as his initial term and the term has an exact match in the CAB thesaurus. There is also an option below the exact match term which is labelled as “search as keyword”. The user can tick the box and ask the system to perform a free-text search on the term.

There is an option in front of the matched terms named “auto explode” which automatically includes narrower terms of the matched term to broaden the search results.

There is an option labelled as “scope” which is scope note for finding more about the application of the term.

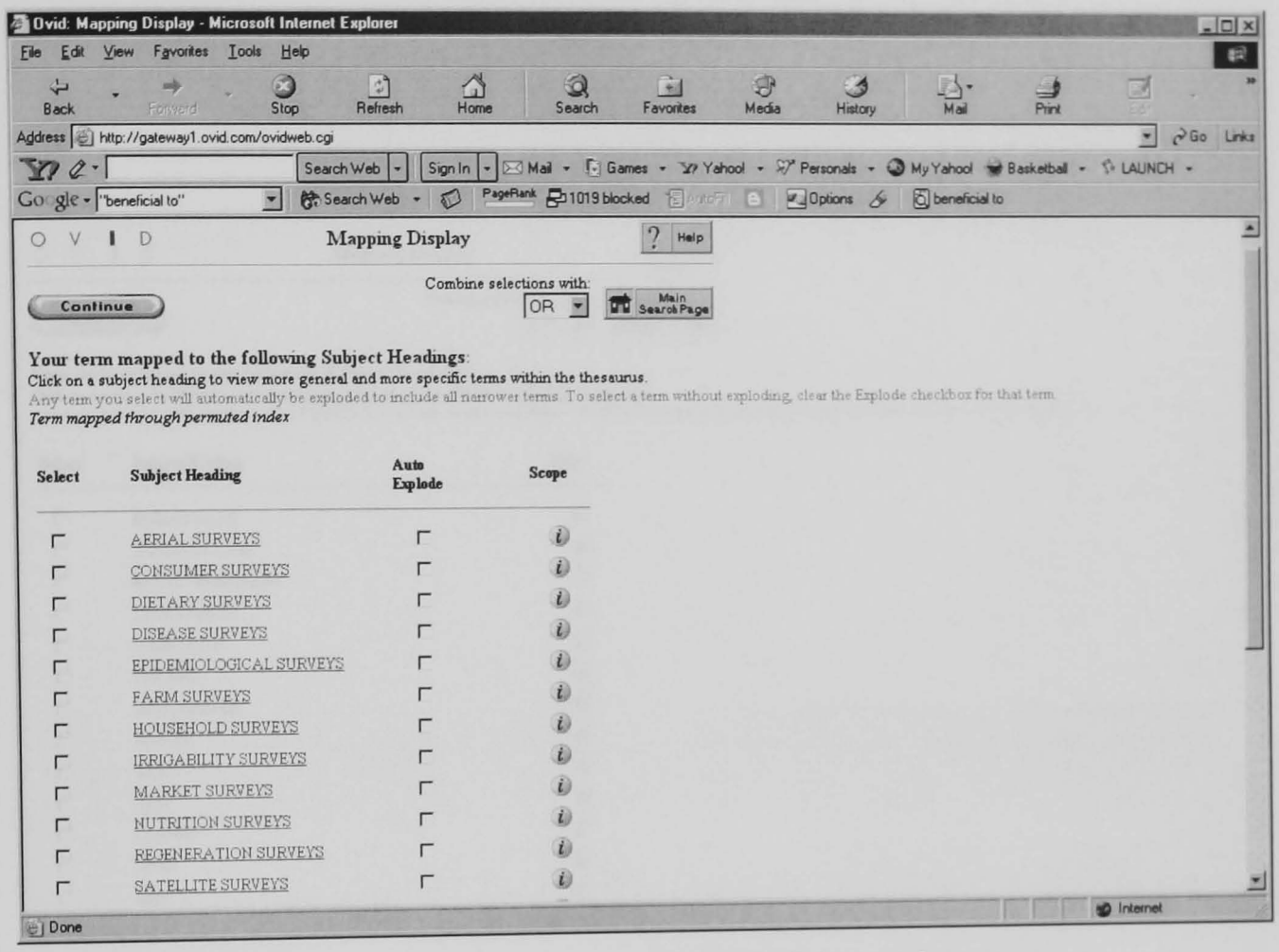


Figure 3. An example of a partial match page

Figure 3 shows a partial match example where the user enters “survey” and the term matches partially with terms that include “survey” such as “farm surveys”, “consumer survey” and so forth.

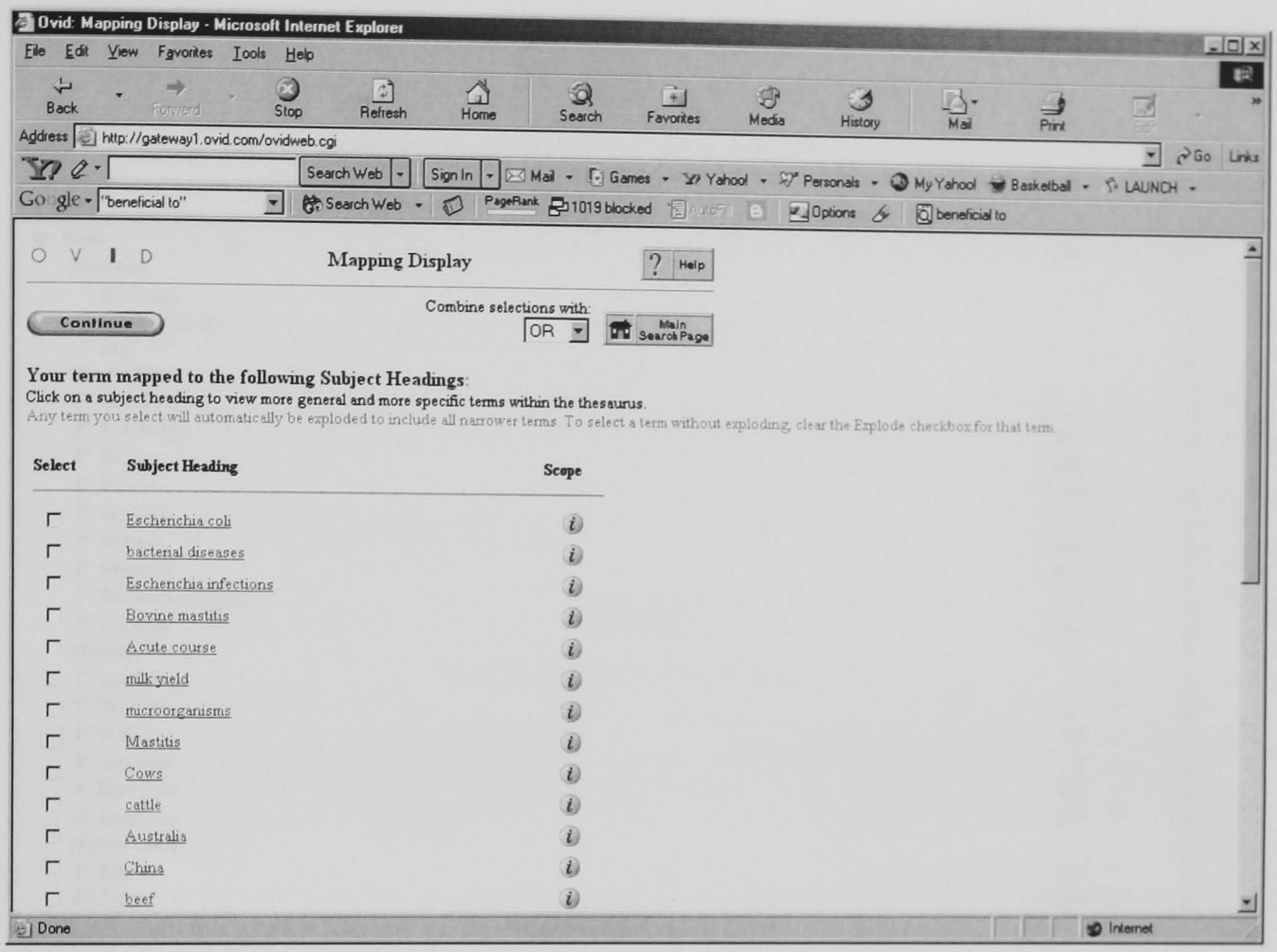


Figure 4. An example of a statistical match page

Figure 4 shows an example of a statistical match where the user has entered the term “Ecoli 0157” and terms like “Australia” and “China” have also been suggested to the user.

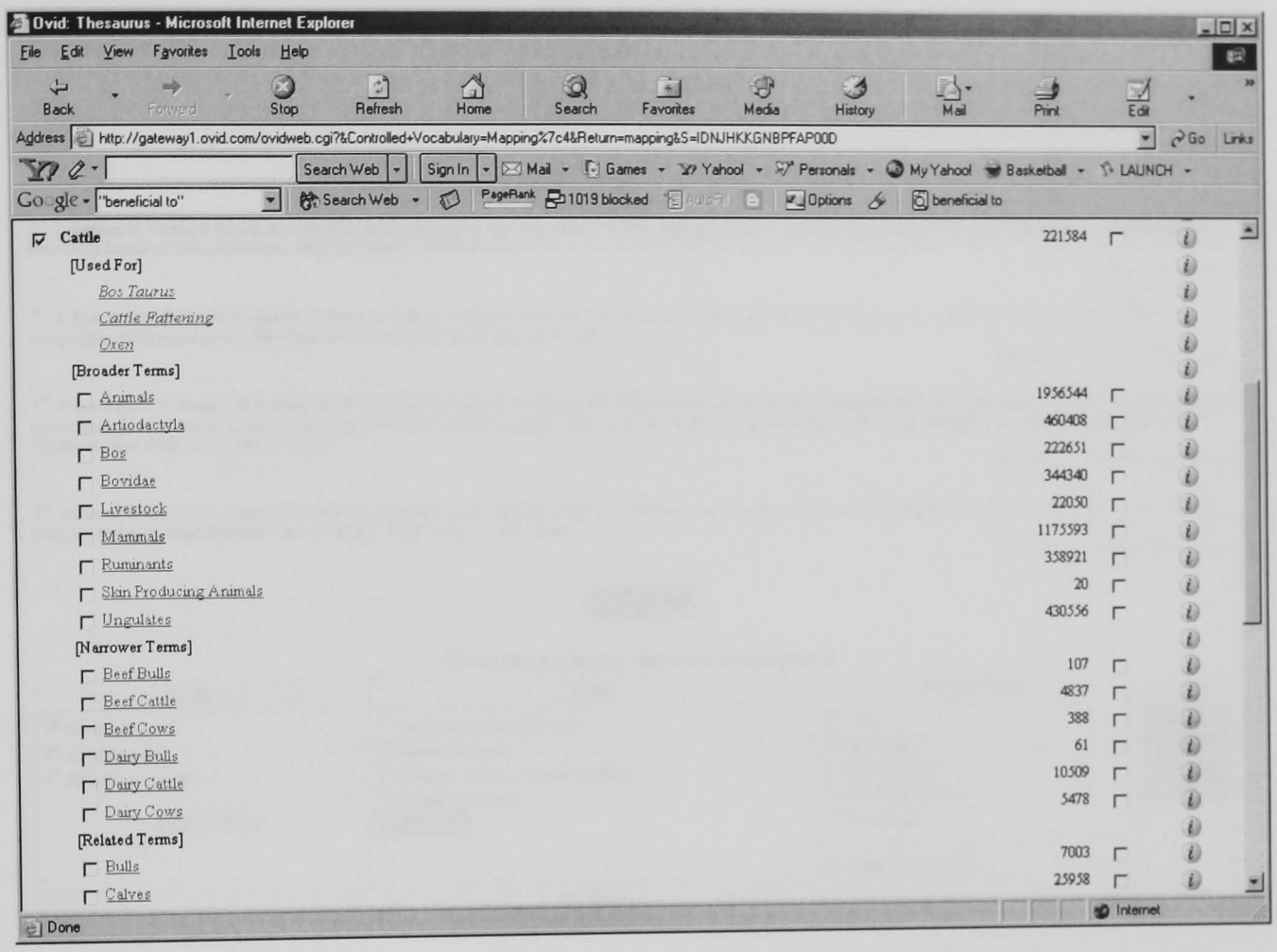


Figure 5. Hierarchical structure of the CAB thesaurus

Figure 5 shows an example of a hierarchical display of the term “cattle” together with all its narrower, broader and related terms. As can be seen, the number of documents indexed by each descriptor has also been indicated.

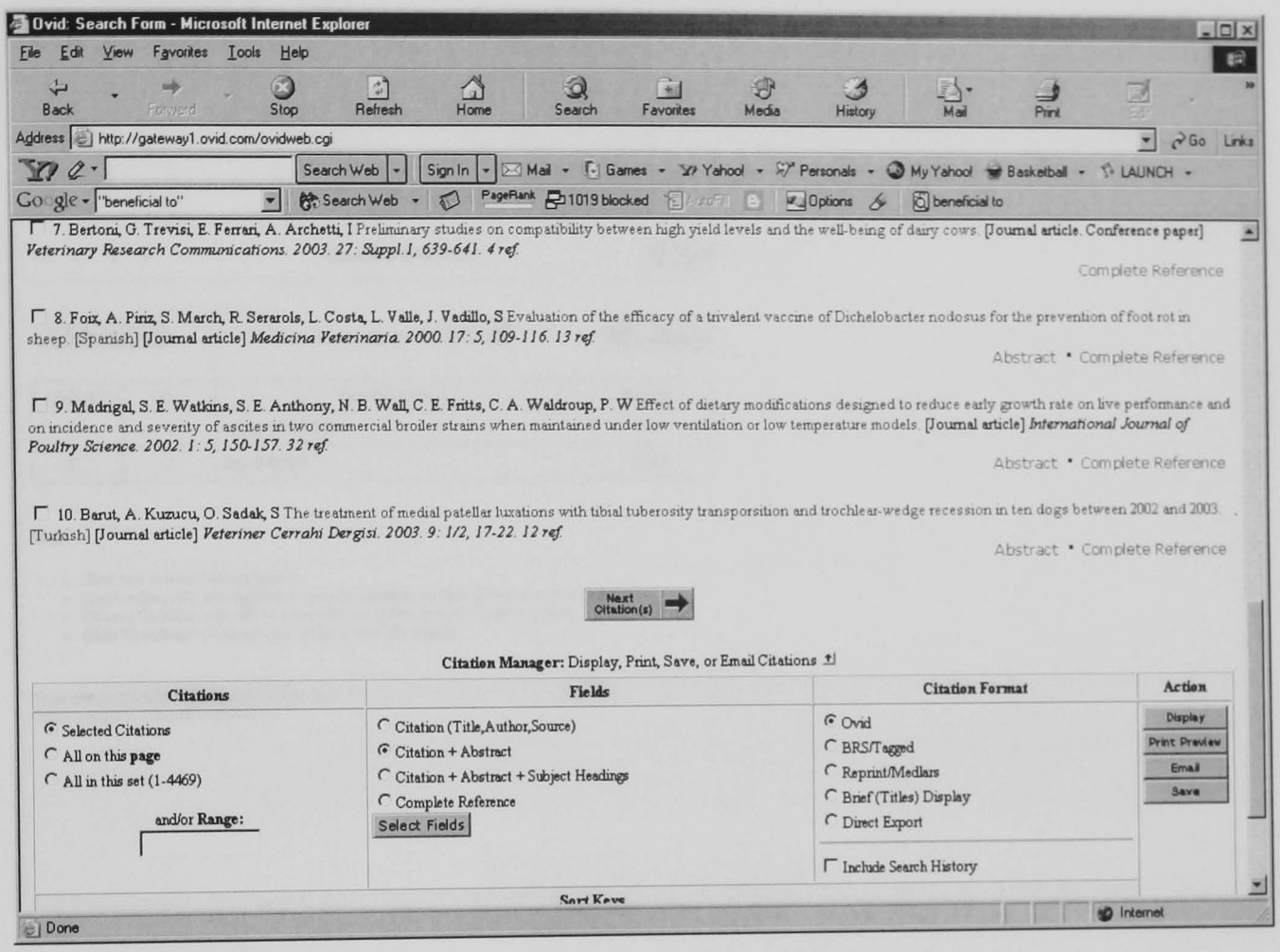


Figure 6. An example of the results display page and citation manager

Figure 6 shows an example of a results display page with the "citation manager" feature whereby users can choose what fields and formats to be included in their result sets.

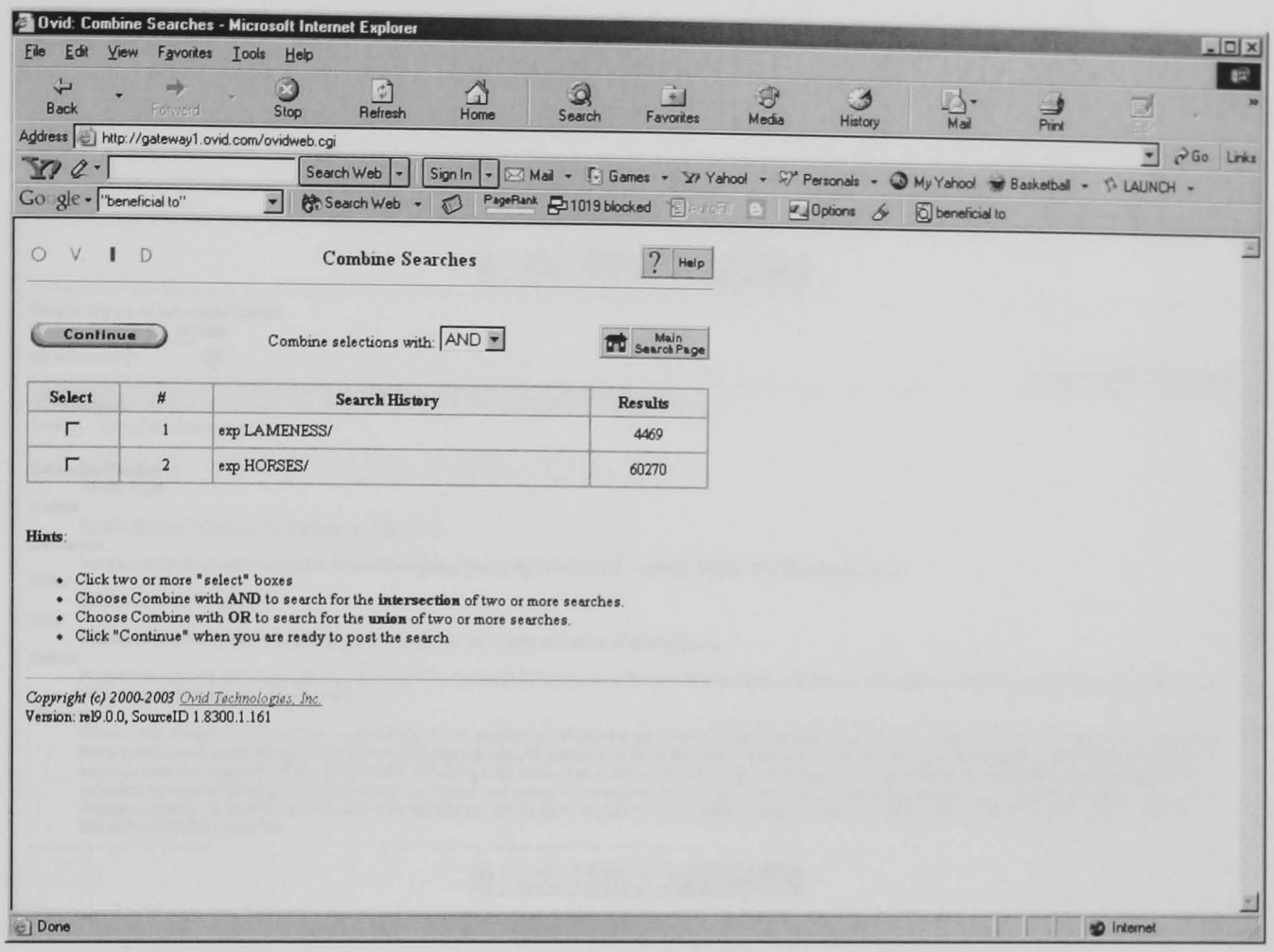


Figure 7. Search combination page

Figure 7 shows an example of a search combination page where the user can combine the two terms “lameness” and “horses” using AND or OR.

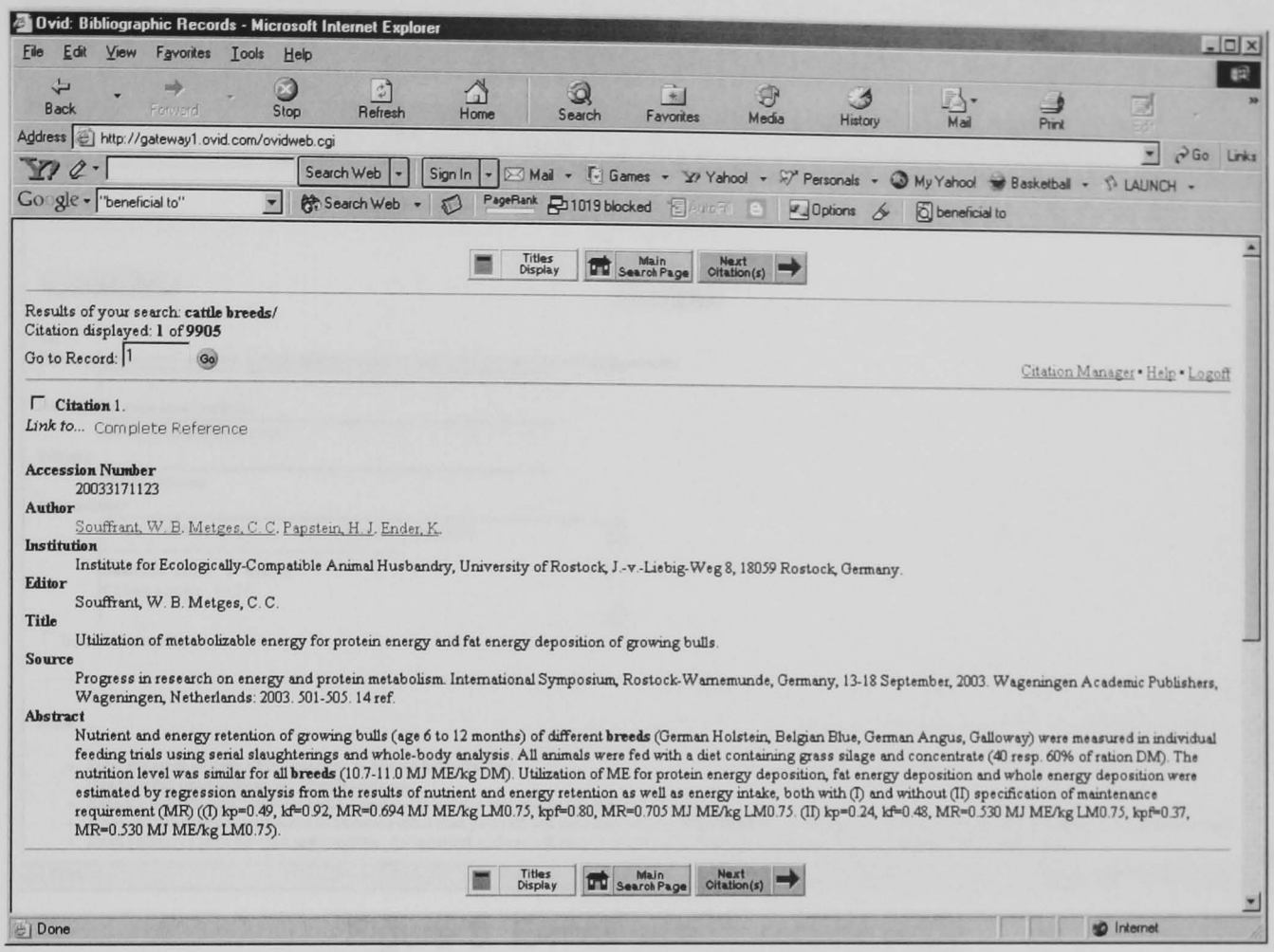


Figure 8. Abstract of a bibliographic reference

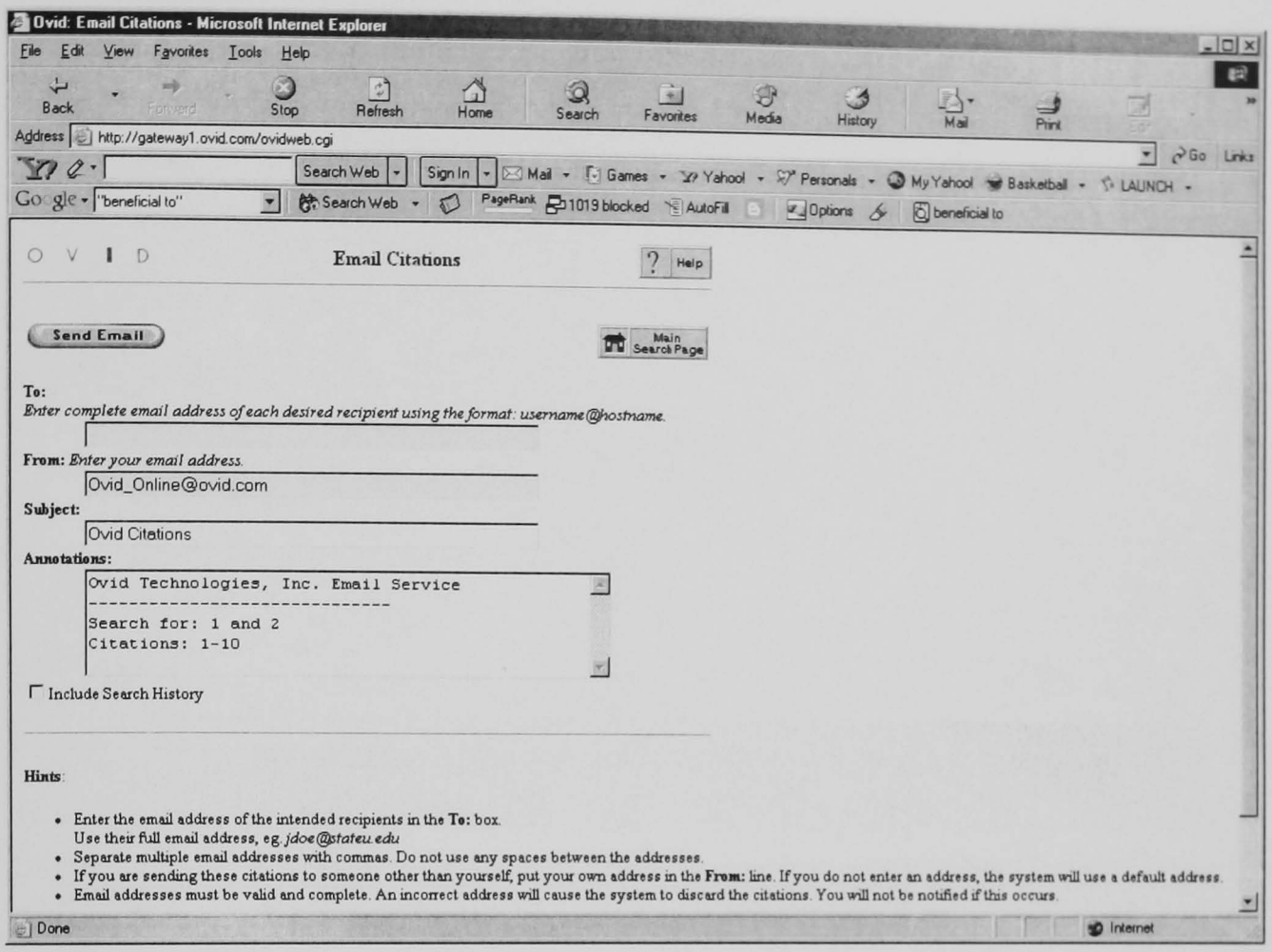


Figure 9. E-mail search results page

Figure 9 shows an example of e-mail feature on the Ovid search interface where users can send their search results through e-mail.