THE IMPACT OF ALTERNATIVE SEARCH MECHANISMS ON THE EFFECTIVENESS OF KNOWLEDGE RETRIEVAL

by

Ryan Craig LaBrie

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

ARIZONA STATE UNIVERSITY

August 2004

UMI Number: 3135262

THE IMPACT OF ALTERNATIVE SEARCH MECHANISMS ON THE

EFFECTIVENESS OF KNOWLEDGE RETRIEVAL

by

Ryan Craig LaBrie

has been approved

May 2004

APPROVED:

_________________________, Chair

____

____

____

____

Supervisory Committee

ACCEPTED:

____

Dean of the School

____

Dean, Graduate College

ABSTRACT

Intangible knowledge assets now account for over 70% of organizational assets. With this has come the need for organizations to invest heavily in knowledge management systems (KMS). For a KMS to be effective, it must be utilized. A critical part of being utilized is having an effective retrieval mechanism. Due to the complexity of knowledge objects stored in KMSs, traditional data retrieval methods utilizing keyword search capabilities may not be the optimal retrieval mechanism for retrieving knowledge. The purpose of this dissertation is to determine if the cognitive loading of a search mechanism impacts the effectiveness of information retrieval from knowledge management systems. To answer that question, two search interface mechanisms are created: one with a text-based keyword mechanism similar to most current search interfaces and one with a visual tree-view hierarchy-based search mechanism. A laboratory study is performed to compare measures of users' accuracy, timeliness, work effort, and satisfaction on those two mechanisms for three different search scenarios.

Theories developed in cognitive psychology based on the recall versus recognition paradigm suggest that the hierarchical nature of the visual tree-view search interface mechanism will generate more accurate information. It is also predicted that the visual tree-view search interface will result in slower searches, but that trade-off will offer more accurate information to the knowledge worker. These predictions lead to directional hypotheses that can be tested in an experimental setting. The results from this experiment show retrieval accuracy is significantly increased by utilizing a visual hierarchical-based search interface. The results also show there is a difference in time and effort but little to no difference in satisfaction between the two search interfaces.

Based on these findings, it is posited that knowledge management systems that are designed with more effective retrieval mechanisms (i.e., visual search interfaces) can reduce organizational costs by increasing relevant information retrieved during the search process and decreasing the wasted time spent investigating information that does not support the current knowledge requirements of the user.

This work is dedicated to my Father.

ACKNOWLEDGEMENTS

Like most published work there was a great deal of encouragement and assistance by many people which helped me to bring this dissertation to its completion. First and foremost I would like to thank my wonderful family. Without their encouragement and support I never would have had the motivation to pursue this dream. To my best friend and wife, Jennifer and my beautiful children Rachel and Alexander, let us pray this is a new beginning where Alexander doesn't have to ask, "Where's Daddy?" so often. To my parents, who instilled in me that hard work pays off, thank you. To my in-laws, thank you for sharing your wisdom and our home four months each year during our time in Phoenix, your time with us was always more of a blessing than a burden.

I would also recognize each of the members of my committee for their individual contributions to this work. Dr. Karen Dowling, you provided guidance and support even before I arrived at ASU. Additionally, your detailed reviews and thought provoking questions made this dissertation stronger. Dr. Haluk Demirkan, your addition brought a fresh perspective and an encouraging comradeship in which I was able to feel, "Yes, I can get through this!" Dr. Andrew Chen, in addition to your committee service, thank you for working with me on my first scholarly paper. While the first conference rejection was hard, the reward of a publication that came from our efforts was well worth it, may many more follow. Dr. Ajay Vinzé, your leadership though the doctoral program has made this experience worthwhile. Thank you for the opportunities in CABIT, may your research center continue its success. Thank you also for making me consider the big picture perspective about my chosen research stream, this too has made my research stronger.

To Dr. Robert St. Louis (lovingly referred to by my children as "Dr. Bob"), you amaze me. When I grow up I hope to be just like you… statistically significant! You are a true scholar, leader, and mentor. Words cannot express the appreciation I have for the many hours you selflessly shared with me and the opportunities you provided.

To Dr. Elaine Weltz, I thanked you in my Master's thesis and I thank you again in my Doctoral thesis. It was in your undergraduate database class in 1988, when I decided what I wanted to do with my life… and it only took 16 years to get there! To Dr. Gerhard Steinke, thank you for your mentorship and nudging me into teaching my first class. To Dr. Alec Hill, thank you for not hiring me full time until I earned my Ph.D. You have made me grow in ways I did not believe possible. To William H. Gates, III, thank you for creating the world's greatest company. The experiences and opportunities I had while at Microsoft were invaluable. Those ten years made this education possible, and will make the education I provide others that much more valuable.

I would also like to take this opportunity to thank two others that have shared in this experience with me, shoulder to shoulder, side by side, Joseph Cazier and Beom-Jin Choi. Gentlemen, we have survived and we have succeeded. I thank each of you for your fellowship, may the bonds we created here at ASU never diminish.

Finally, to my golf foursome, Andrew, Craig, and Ray, thank you fore! our times together. An important factor in any life's pursuit is a work-play balance. While these four years haven't been very balanced, I appreciated each time we were able to get together. I hope that we will continue our friendship and our foursome in the years to come, no matter where we find ourselves.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Motivation and Background

Within the information systems (IS) community, both from an academic and industry perspective, there are significant levels of attention being paid to the topic generally referred to as Knowledge Management (KM). It is reported that in 2002, U.S. companies spent $4.5 billion on knowledge management solutions (Gilmour, 2003). The U.S. Federal Government spent $820 million in fiscal year 2003 on knowledge management initiatives and intends to increase that spending at a 9% compound growth rate to nearly $1.3 billion by 2008 (Business Wire, 2003). A recent query[1] on the Internet search website, GOOGLE.COM, produced 3,150,000 web pages related to "knowledge management." Searching for books on "knowledge management" produced 1892 results from the online bookstore AMAZON.COM during that same period. Furthermore, business school academics have been contributing to this widely investigated topic of knowledge management by producing courses, lectures, books, scholarly journals, and research articles dedicated to knowledge management and knowledge management systems (KMS).

Research conducted under the auspices of knowledge management varies greatly in direction and scope. In fact in recent years, entire issues of leading information systems research journals have been dedicated to the topic of knowledge management and knowledge management systems. Examples include special issues from *Journal of*

---

[1] Searches where performed on 1 November 2003.

Management Information Systems – volume 18: issue 1 in 2001, *Management Information Systems Quarterly* – volume 26: issue 3 in 2002 and again another *MISQ* issue forthcoming in 2004.

Alavi & Leidner (2001) provide a foundational framework to help contextualize the various components of knowledge management systems research and development. In this framework they offer four focus areas for research; knowledge creation, knowledge storage/retrieval, knowledge transfer, and knowledge application. This dissertation work, while recognizing the need for research in all four areas of knowledge management domains, focuses specifically on the retrieval aspects from knowledge management systems. Most information system professionals are aware of the acronym GIGO – garbage in, garbage out. In fact, for well over 30 years there has been a significant amount of research that has focused on developing information systems to decrease the garbage-in factor. This research focuses on the problem where you might have quality information captured in a knowledge management system, but do not have an effective mechanism to retrieve that knowledge. Thus you are faced with the dilemma of quality in, but still garbage out.

In her theory of knowledge reuse Markus (2001) suggests that the information system that supports the knowledge management initiative should handle both access to expertise and access to the experts. This allows for optimally managing both the explicit knowledge of the organization, the knowledge that is captured in the system as expertise, as well as an attempt to manage the tacit knowledge, which resides within the individuals (experts) of the organization. While this dissertation research focuses primarily on explicit knowledge that has already been captured and codified in electronic format, its implications may extend beyond expertise and apply to knowledge management of experts as well by helping to manage individuals with implicit knowledge. In developing this theory of knowledge reuse Markus (2001) stresses that information technology plays an intermediary role. Specifically, Markus places greater emphasis on the creation and utilization of the repository. This dissertation echoes that emphasis and suggests that by re-evaluating the underlying infrastructure of the repository, the information retrieval process may be more effective.

Research Question

Fundamentally, this dissertation focuses on the following research question:

Does the cognitive loading of search mechanisms impact the effectiveness of knowledge retrieval?

For this study, cognitive loading refers to a recall versus recognition paradigm; retrieval mechanisms are operationalized as a keyword search method or a visual tree-view hierarchical search method; and effectiveness is operationalized as measures of accuracy, timeliness, work effort, and satisfaction.

To answer this question an experiment was designed to test two different search methods, one based on a traditional text-based keyword search interface and one based on a visual tree-view hierarchical interface. Both systems utilized the same data set. Hypotheses based on prior theory are proposed, the experiment is discussed, and results are analyzed and presented.

The contributions of this research are multifold. First, it builds on prior theory. A purposeful attempt is made to incorporate theory from the cognitive psychology domain; specifically findings in recollection versus recognition (Simon, 2001; Anderson, 1995; Bower et al., 1969). What we have learned about human nature, learning, and knowledge use from a psychological perspective needs to be considered when developing information systems for knowledge (re)use. Information systems research could benefit greatly if it would draw more often from more mature reference disciplines, especially when it relates to human-computer interactions and knowledge. This study is a single attempt to add to that limitation. Second, this research seeks to address a problem that continues to haunt users of information systems – What is the best way to extract knowledge from information systems given the issue of information overload? Technological advances in file/storage management, indexing and wildcard capabilities have taken us a long way in finding the information we need. However, without a fundamental shift in our thinking in designing these knowledge management systems, our retrieval efficiencies may be reaching a plateau. Information overload is a real problem that continues to get exponentially worse with each passing year. Case in point, it took 100,000 years to create the first 12 exabytes of information, the second 12 exabytes was created in 2.5 years, and the third 12 exabytes of data will be created in 1 year (Ball, 2002). Figure 1 shows the development of information, some common milestones, data definitions, as well as human capacity and world population.

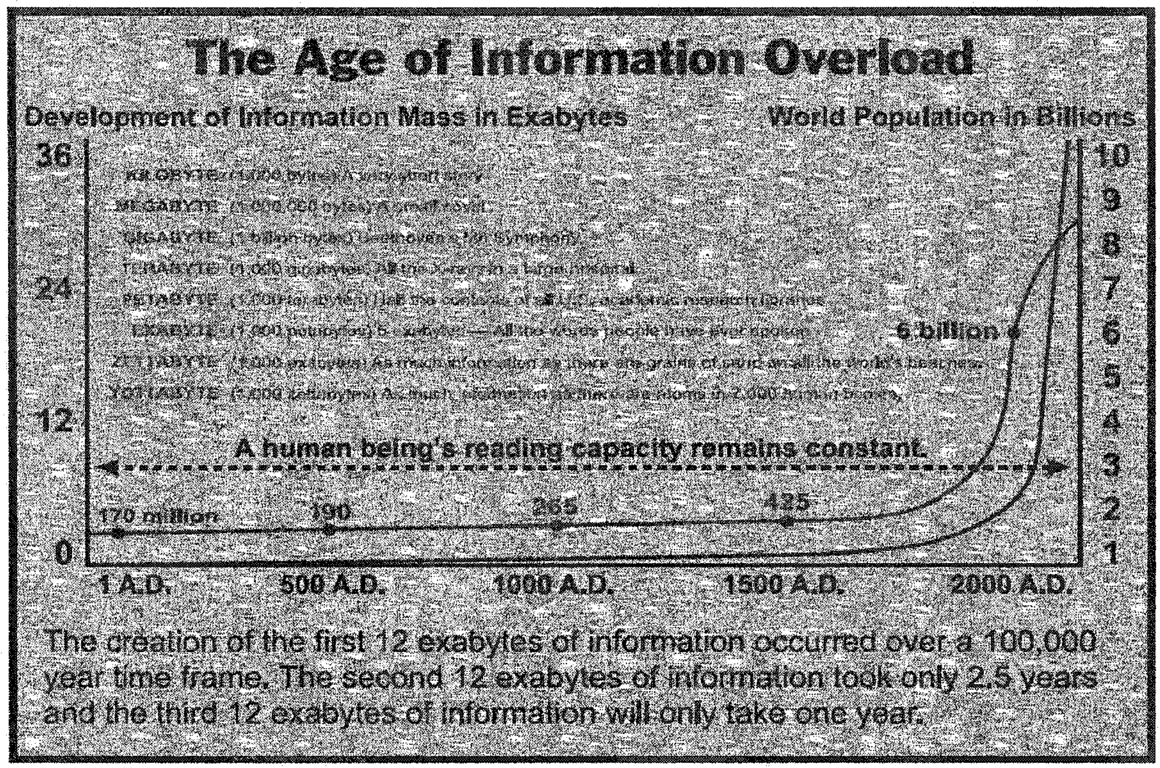


Figure 1. The history of information creation (source: Ball, 2002).

Third, there is little literature in this research domain; in fact, this research is a direct response to Alavi and Leidner's (2001) call for research in knowledge management, specific to storage and retrieval. While much of the research on knowledge management has been qualitative in nature this research brings an experimental approach with empirical evidence that tests *a priori* hypotheses. Finally, there are important implications to both academics and industry that can be gathered from the results of this research. The findings are expected to indicate that knowledge management system developers need to make radical departures from their typical retrieval user interfaces to increase accuracy of KMS use as hypothesized in this research, or to deal with sub-optimal retrieval results.

Organization of Dissertation

The remainder of this dissertation is organized as follows. Chapter 2 reviews prior literature; this review provides foundational and theoretical contexts in which this research is conducted. Knowledge management literature is first examined, beginning with definitions of data, information, and knowledge, moving through a discussion of ontology, taxonomy, and controlled vocabulary. Next, a discussion on the contributions of document management and information retrieval is presented. The retrieval research is critical to this work as it is posited that searching and retrieving knowledge is different than searching and retrieving data. An examination of the limitation of keywords as the primary means for search and classification alternatives is then discussed. The literature review then concludes with a review of the recall versus recognition paradigm literature from the cognitive psychology domain. In Chapter 3, a model is developed and testable hypotheses are proposed based on the knowledge gained from the theoretical grounding from prior research. Chapter 4 presents the methodology used to conduct the research for this dissertation. In this case, a laboratory experiment is developed and administered to collect the data used for testing the hypotheses. Within this chapter experimental design and statistical analysis methods used to evaluate the data are also discussed. Analysis of the results from the data collected during the experiment is presented in Chapter 5.

Chapter 6 concludes this dissertation with some discussion of the implications of the findings, presents some study limitation, and provides avenues for future research.

CHAPTER 2

LITERATURE REVIEW

The foundations for this research are developed through the following bodies of literature. First, an investigation of current knowledge management is presented. The knowledge management research highlights what knowledge management is and where knowledge management research should be focused via the presentation of a number of research frameworks. As a foundation to understanding knowledge management literature is reviewed that documents there the relationships between data, information, and knowledge. Another important foundation for knowledge management research is an understanding of ontologies, taxonomies, and controlled vocabularies. Next information retrieval literature is examined. This critical body of research helps to define and support the differences between searching and retrieving data in its most basic forms of character text and numbers through multimedia based data such as images and voice, to the knowledge objects requested today in more advanced knowledge management systems. In addition to the implications of data search versus knowledge search, the limitations of keywords and classification are examined as mechanisms for search and retrieval. Included in the information retrieval section is discussion of the electronic document management systems literature. The literature represents an early example of knowledge management systems. Representing the foundational underlying theories of this research, a review of the recall versus recognition literature from the cognitive psychology domain is examined. The literature review closes with discussion of works that have focused on effort versus accuracy within information systems.

Knowledge Management and its Research Foundations

Knowledge management research is a very broad area within the information systems discipline. Several authors (Alavi & Leidner 1999, 2001; Spiegler, 2000; Schultze & Leidner, 2002) have suggested various classification models that break down knowledge management into different domains. Alavi and Leidner (1999) surveyed 109 executives, obtaining 50 usable responses, on their perceptions of KMS activity within their firms and its potential benefits. This research identified three perspectives for knowledge management – an information-based, a technology-based, and a cultural-based perspective. Among the conclusions from this research were: 1) knowledge management systems are multi-faceted; and 2) it is important to develop metrics to assess the benefits of KMS. In their 2001 work Alavi and Leidner (2001) propose a conceptual foundation that includes the knowledge management systems domains of knowledge creation, knowledge storage/retrieval, knowledge transfer, and knowledge application.

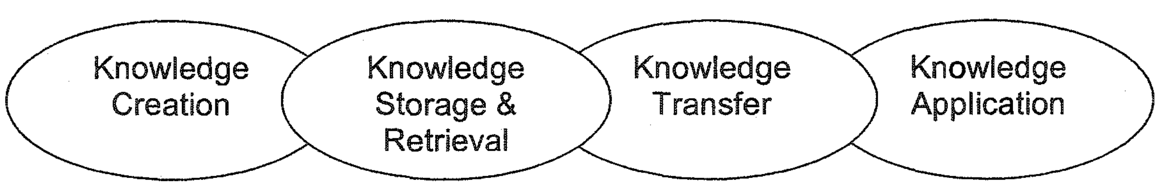


Figure 2. A knowledge management research framework.

Strategic research questions posed by Alavi and Leidner (2001) pertaining to knowledge retrieval includes:

1. Is stored knowledge accessed and applied by individuals who do not know the originator of the knowledge?
2. What retrieval mechanisms are most effective in enabling knowledge retrieval?

Some have questioned whether or not knowledge management research is actually a new idea. In his work Spiegler (2000) concludes that knowledge management is indeed a new idea, rather than a recycled concept. The basis of his argument is that yesterday's data are today's information, and tomorrow's knowledge, which in turn, recycles back through the value chain of data-information-knowledge, represents another framework in which to investigate knowledge management.

Schultze and Leidner (2002) propose using Deetz's four discourses of organizational inquiry to classify IS knowledge management research. These categories include: normative, interpretive, critical, and dialogic discourse. Based on an analysis of the literature they find that the bulk of the research falls within the normative discourse classification, interpretive discourse is used less frequently, and lagging relatively far behind is KM research following a critical or dialogic discourse. Similar to the Alavi and Leidner (2001), this classification of knowledge management research helps to provide a context in which knowledge management is studied from the IS discipline's perspective.

Each of the articles discussed above suggests various frameworks or classifications for knowledge management research within the information systems field. Table 1 summarizes their contributions.

Table 1

Knowledge Management Research Classifications

Author (Year)	KM Classification Context	Research Method
Alavi & Leidner (1999)	KM Perspectives: Information-based, Technology-based, and Culture-based	Executive Survey (109 Surveyed, 50 usable responses)
Alavi & Leidner (2001)	KM Processes: Creation, Storage/Retrieval, Transfer, and Application	Literature Review
Spiegler (2000)	KM Transformation: Data Processing, Information Processing, Knowledge Processing, and a reverse process	Literature Review and Model Development
Schultze & Leidner (2002)	KM Discourses Normative, Interpretive, Critical, and Dialogic	Literature Review

Data, Information, and Knowledge.

To understand the functions of knowledge management systems (KMS) it is critical to differentiate them from database management systems (DBMS). Fundamentally the difference boils down to the content being managed. In database systems typically what is stored and manipulated are atomic items, usually in the form of numeric or text representations. In contrast, knowledge management systems store and manipulate complex, *unstructured* objects. Objects here refer to electronic items that can vary greatly in ability and content; for example a knowledge object may be a word processing document, an electronic spreadsheet, a multimedia video clip, etc. These objects typically additionally maintain their own set of metadata that also must be managed and that can be utilized in the search and retrieval of knowledge objects. The management of these types of objects cause for reevaluation of traditional ways of managing electronic data.

One of the major challenges knowledge management researchers have is a lack of standard definitions of some of its most basic constructs. Most researchers agree that when dealing with knowledge management one must have an understanding of what data, information, and knowledge are and what the relationships between them are, however there are no clear standards for those definitions (Moore, 2002; Hick et al., 2002; Meadow & Yuan, 1997; Teskey, 1989). Most agree that data are "raw facts" observed from some type of event. Additionally, most definitions of information include the idea that it is "structured data" or "data within a context". Where knowledge management researchers have trouble is with the definition of knowledge. While it would be convenient to follow Raisinghani's (2000) formulation of data, information, and knowledge definitions:

- Data are raw facts;
- Information is formatted data;
- Knowledge is formatted information;

These definitions should be considered over-simplified. However, what they do suggest is a hierarchical relationship between data, information, and knowledge. This understanding is almost universally accepted by the knowledge management community, even if the actual definitions are not. Table 2 provides a number of definitions on data, information, and knowledge from the knowledge management, computer science, information sciences domains, as well as a traditional dictionary definition. Interestingly, some authors (Van Beveren, 2002; Meadow & Yuan, 1997) suggest that knowledge cannot reside outside the human brain, and as such they purport that knowledge management and knowledge management systems are really misnomers.

Table 2

Definitions of Data, Information, and Knowledge

Source	Definition of Data	Definition of Information	Definition of Knowledge
Moore (2002)	Data corresponds to the bits (zeroes and ones) that comprise a digital entity.	Information corresponds to any tag associated with the bits. The tags are treated as attributes that provide semantic meaning to the bits.	Knowledge corresponds to any relationship that is defined between information attributes. The types of relationships are closely tied to the data model used to define a digital entity. At a minimum, semantic/logical, spatial/structural, temporal/procedural, and systemic/epistemological relationships can be defined between attributes associated with a digital entity, and between the attribute values.
Hicks et al. (2002)	Data is usually considered to be textual, either numeric or alphabetic. Data is considered to be structured and represent a measure such as a quantity.	Describing a fact; which is an occurance of a measure or inference of some quantity or quality. Meaning + measure.	Knowledge is made up of a knowledge element and a knowledge process. The knowledge process is the procedure utilized by the individual to infer the knowledge element from information, other knowledge elements or a combination of each.
Raisinghani (2000)	Data are raw facts	Information is formatted data	Knowledge is formatted information
Meadow & Yuan (1997)	Data usually means a set of symbols with little or no meaning to a recipient	Information is a set of symbols that does have meaning or significance to their recipient	Knowledge is the accumulation and integration of information received and processed by a recipient
Van Beveren (2002)	Data are raw facts	Information is data within a context	Knowledge is an individual's stock of information, skills, experience, beliefs and memories.
Teskey (1989)	Data as the direct result of observation of events	Information as structured collections of data	Knowledge as models of the world, which can be created or modified by new information.
Webster's Dictionary	factual information (as measurements or statistics) used as a basis for reasoning, discussion, or calculation	the communication or reception of knowledge or intelligence	the fact or condition of knowing something with familiarity gained through experience or association

Continuing with the hierarchical relationship of data, information, and knowledge Hick et al. (2002) developed the bi-directional decision-making model presented in Figure 3. One of the more important aspects of this model is that information and knowledge both lead to decisions as well as decisions producing additional information and knowledge. Additionally, there is a recursive relationship between information and knowledge, where information feeds into knowledge – as in the simplified definition from Raisinghani (2000), and knowledge feeds back into information – a factor that is missing from Raisinghani's definitions.

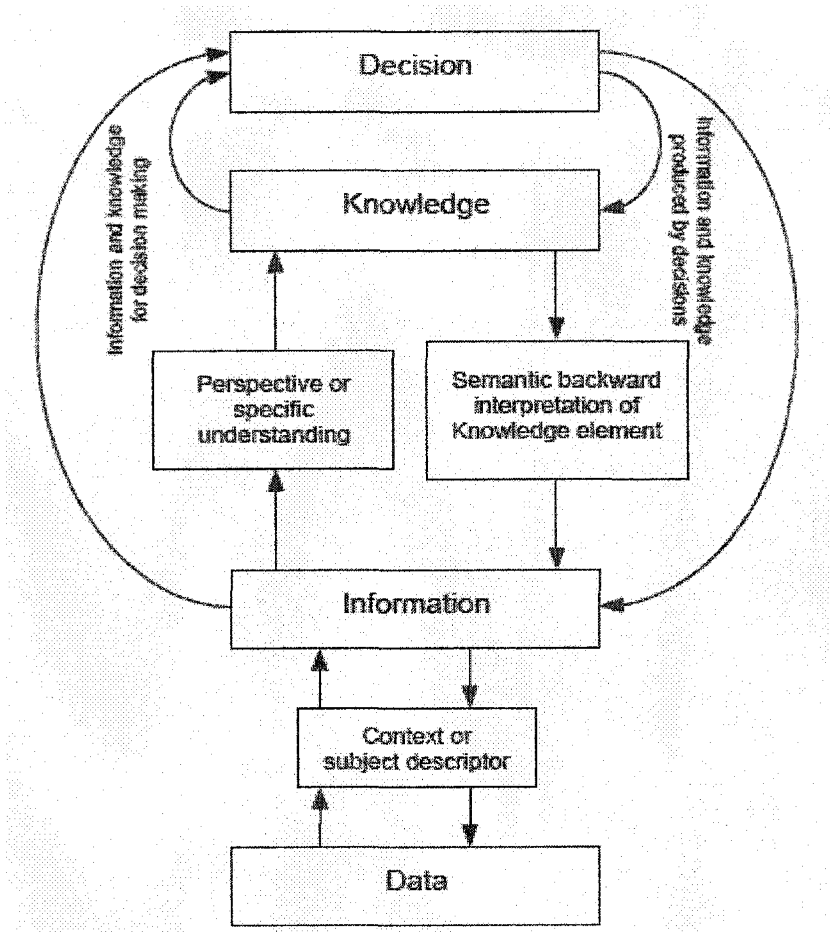


Figure 3. A bi-directional decision making model (Source: Hicks et al., 2002).

Ontologies, Taxonomies, and Controlled Vocabularies.

Another issue the knowledge management community is currently wrestling with is the understanding and integration of ontology in KM research (Chandrasekaran et al., 1999; Swartout & Tate, 1999; Edgington et al., in press). Ontology is defined as the basic structure around which knowledge can be built. An ontology provides a set of concepts and terms for describing some domain, while a knowledge base uses those terms to represent what is true about some real or hypothetical world (Swartout & Tate, 1999).

Ontologies differ from taxonomies in that ontologies can be analogous to, in the database world, entities and their relationships, while taxonomies can be likened to specific instances of an entity, that is, the concrete manifestation of reality (i.e. a table, directory, or tree structure). From there, controlled vocabularies specify the ontological domain's keywords and thesaurus. Edgington et al. (2004) capture this in their ontology creation process, shown in Figure 4.

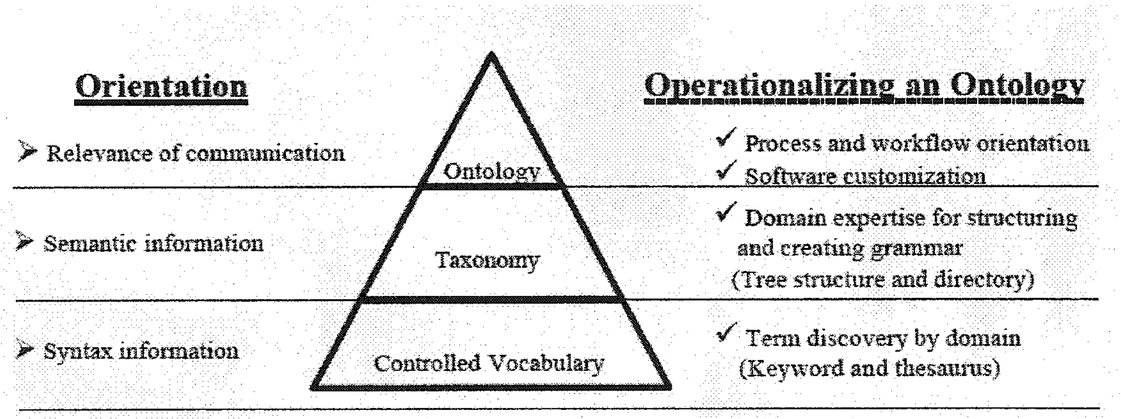


Figure 4. The controlled vocabulary, taxonomy, and ontology hierarchy (Source: Edgington et al., 2004)

As with data, information, and knowledge, a similar hierarchical relationship exists between controlled vocabulary, taxonomy, and ontology. The use of ontologies within the IS/KM domain has been examined (Hori, 2000; Wand and Weber, 1990, Wand et al., 2000; Edgington et al., in press) and has mainly focused on the activity of a knowledge sharing process. However, not all are in agreement as to the use of ontologies as a panacea for the knowledge sharing process (Correa da Silva et al., 2002). Their research highlights a number of shortcomings with regards to ontologies, including the difficulty of reuse, and the difficulty of sharing inference, semantic, and group knowledge. Implications of ontological research within the knowledge management domain provided structure on dealing with the cognitive overload that knowledge objects can cause in an information overloaded environment.

Information Retrieval

As a precursor to the knowledge management literature, some pertinent insights can be gained from the information retrieval literature and the document management literature. Document management systems have existed for a number of years and can arguably be considered some of the earliest forms of knowledge management systems. Sprague (1995) describes how information systems managers, if properly prepared, can take the next step beyond managing text and numbers to managing electronic document objects. These objects may take the form of contracts, email/voicemail, video clips, meeting transcripts, drawings/blueprints/photographs, or any number of object types. The contributions of this research include the idea that managing knowledge objects is different from managing basic text and numbers. Additionally, Gordon and Moore (1999) develop a foundation for a "readying system" that examines how a document is used and the purpose for which it is used. This readying system is a new type of information system developed with a formal language to help knowledge workers retrieve knowledge documents in a more effective manner.

One of the primary research streams within the information retrieval literature is document management. Several researchers have focused specifically on information retrieval within a document management perspective (Gordon, 1997; Blair, 2002a, 2002b; Blair & Kimbrough, 2002). Similar to the ontological based research in knowledge management which focuses on knowledge sharing, the information retrieval research largely focuses on knowledge sharing by improving methods of search. The information retrieval community largely agrees that providing full text search and/or faster better ways for indexing is not the solution to the business world's problem of information overload (Moore et al., 1990; Blair & Kimbrough, 2002; Blair, 2002a, 2002b). Information retrieval researchers are constantly trying to find the new and improved search and retrieval mechanisms. As an example, exemplary documents have been suggested as templates of best practices from which to develop indexing schemes (Blair & Kimbrough, 2002).

Blair (2002b) draws a distinction between retrieving data and retrieving documents. His definition of document retrieval actually can be directly applied to a knowledge management environment and KMS retrieval. Fundamentally, the differences include: direct vs. indirect searches, different success criterion, and different speed dependencies. Table 3 identifies five major distinctions between data and document (knowledge) retrieval.

Table 3

Differences between Data and Knowledge Retrieval Systems (Source: Blair, 2002a)

Data Retrieval	Knowledge (Document) Retrieval
Direct ("I want to know X")	Indirect ("I want to know *about* X")
Necessary relation between a formal query and the representation of a satisfactory answer	*Probabilistic relation* between a formal query and the representation of a satisfactory answer
Criterion of success = correctness	Criterion of success = utility
Speed dependent on the time of physical access	Speed dependent on the number of logical decisions the searcher must make
Scales easily	Does not scale easily

The distinction between "direct" and "indirect" is in direct retrieval the correct answer is there and you know it, whereas in indirect retrieval a correct answer may or may not exist. The distinction between the criteria for success differs as well between a data retrieval system and a knowledge retrieval system. In a data retrieval system success is related to correctness, that is, the correct answer(s) is (are) retrieved and there is no ambiguity. In successful knowledge retrieval experiences the system retrieves the most relevant knowledge object(s). Relevance is thus determined via utility which is a much more ambiguous measure of success. In data retrieval environments speed can be controlled via faster hardware and software techniques. In knowledge retrieval environments speed is less related to the hardware and software, but rather more related to the knowledge workers trial and error attempts and number of iterations they utilize in their search effort. Finally, there is the issue of scaling. In data retrieval systems hardware and software can be used to scale the system easily. Going from 100 to one million data records is relatively easy. Moving up to one billion, or even one trillion, data records can be handled with simple additions of appropriate hardware. However, moving from 100 to one million knowledge objects can have highly negative impacts on knowledge retrieval. Search and retrieval strategies must be adjusted for the knowledge worker in order for knowledge retrieval to be effective.

Keywords versus Classifications.

An additional set of literature that is examined pertains to codification/ classification systems and keyword usage (the term keyword is used in this study to mean both keyword and keyword phrases). Keywords play an important role in information retrieval, yet they have their shortcomings as well. These shortcomings include; ambiguity, forced recollection, user's spelling ability, and a lack of controlled vocabulary to name a few. Often to overcome these limitations classification categories are developed. Perhaps the most widely recognized work on keyword classification systems within the IS community is the work of Barki et al. (1988, 1993). They have developed a classification system commonly referred to as the ISRL (Information Systems Research Literature) categories. These categories were developed based on keyword usage in top IS journals and is utilized by leading IS journals including *MIS Quarterly*.

Similar to the methods used by Barki et al. (1988, 1993) other IS researchers have used analysis of publications to develop classification schemes (Neufeld & Staples, 2002; Vessey et al., 2002; Swanson & Ramiller, 1993; Gorla & Walker, 1998). Vessey et al. (2002) analyzed the diversity within the IS discipline and its journals. They produced a classification scheme based on the following five categories: reference discipline, level of analysis, topic, research approach, and research method. Similarly, Swanson and Ramiller's (1993) paper on information systems research thematics analyzed submissions to a new journal to discover themes and relationships among IS research. Neufeld et al. (2002) explored the relationship of IS topics published in non-IS business disciplines (such as accounting, marketing, etc.).

Gorla and Walker (1998) suggest that searches can not be effective unless an unambiguous keyword list is universally accepted. Findings from their research show the frequency of keyword use is low primarily due to ambiguities such as:

- Errors in spelling;
- Inconsistencies of abbreviations;
- Improper combining of no-similar terms;
- Inconsistent spelling of words;
- Inconsistent compounding of words;
- Redundant keywords.

This leads to an unnecessarily faster growing domain of keywords that ends up making information retrieval more difficult. Gorla and Walker collected their data from the ABI/Inform database for the top MIS journals from 1984-1994. Their analysis was performed on 14,676 articles, 3305 keywords, and 121,548 occurrences of those keywords.

Similarly, LaBrie and St. Louis (2003) found severe limitation with keyword and keyword searching. Their study examined 1791 keyword used in 608 articles over a 27 year period for the IS journal *MIS Quarterly* producing 2885 keyword-article relationships. While keywords are generally thought to be common identifiers to relate similar documents the finding from their study showed that 77% of all keywords were used only once, and another 13% of keywords were used only twice. This left only 10% of the keywords that were used three or more times. These results help to illustrate the difficulty in searching for articles based on keywords.

One of the largest contributing factors to the problems with keywords is that authors are allowed to choose any words they want as keyword phrases. This leads to all sorts of ambiguous results. As an example, take the case of somebody looking for articles on "system design" in *MIS Quarterly*. They would retrieve three articles. Are there really only three articles published in *MIS Quarterly* about system design? A closer inspection of the data suggests that is far from the truth. Table 4 shows several variations of a "system design" query and the number of articles with that keyword phrase.

Table 4

Keyword Query of "System Design"

Keyword Phrase	Number of Articles
System Design (original query)	3
Systems Design	10
System Designs	0
Systems Designs	1
System Analysis and Design	0
Systems Analysis and Design	4
Information System Design	2
Information Systems Design	8
MIS Systems Design	1
Participative System Design	1
System Design Methods	1
Expert Systems Design	1
Impact and Socio-Technical Systems Design	1
TOTAL	34

Table 4 shows over a 10-fold increase in the possible number of articles about System Design. While it may be argued that some of these are more restrictive versions of system design (i.e. Expert Systems Design or Participative System Design), a strong case can be made that there is no difference between:

- System Design;
- Systems Design;
- System Designs;
- Systems Designs.

Yet depending on how the query is stated the results may be zero (system designs) or 10 (systems design).

Similarly, plurals and abbreviations are also causes for knowledge workers not to be able to retrieve accurate information. Take another popular topic from the information systems research field, decision support systems. In this same data set articles were found that had keyword phrases of:

- Decision Support System;
- Decision Support Systems;
- DSS;
- Decision Support Systems (DSS);
- Decision Support.

While it might be argued that decision support varies significantly from decision support systems, the case could be made that there is no fundamental difference between decision support system, decision support systems, decision support systems (DSS), and DSS.

This issue is one of the primary problems that ontology research is directly attempting to address. Furthermore, intelligent search research, largely being explored in the computer science and engineering disciplines, is making incremental advances in the area of information retrieval. For example, full-text searches now allow knowledge workers to search abstracts or entire bodies of text rather than limiting their querying capabilities to just the keywords provided by the author(s). Furthermore, advances in search algorithms including the integration of artificial intelligence (AI) and fuzzy logic can help. As an example Top-k selection queries can be used to find values without requiring exact matches (Chaudhuri and Gravano, 1999; Chen and Ling, 2002). While recognizing these advances, it seems that much of the business and academic world has yet to incorporate these advances into their knowledge management systems.

As mentioned previously, classification schemes such as the ACM Classification Codes and *MIS Quarterly's* ISRL have been developed, theoretically, to address various keyword limitations. By forcing authors to adhere to a controlled vocabulary, one in which keywords must be selected from a predetermined list; many of the ambiguity barriers of keyword searching are alleviated. This, in theory, should produce better search results.

In that same study, LaBrie and St. Louis (2003) also measured classification scheme usage. The findings of which were very interesting. The analysis of classification codes mirrored those of the keyword findings quite closely. The data set employed 613 unique categorization classifications, occurring 2287 times over 587 articles (21 articles did not provide classification categories). Articles that did not provide any classification categories were typically: issues and opinions, research notes, or introductions to special issues. Like keywords, classification categories are meant to create similar groupings – in this context that would equate to grouping of similar articles. Upon a close inspection of the classification categories used on the *MIS Quarterly* articles we make two observations: 1) a very small set are used over and over again to classify the article in the most generic sense, and 2) like keywords, the bulk of the classifications are used a relatively small numbers of times. 45% of all classification codes were used only once, another 18% of the classification codes where used only by two articles, leaving 37% of the classification codes to be used three or more times.

It appears based on this analysis that classification codes are doing little to help solve keyword limitations. Classification codes have their own set of limitations. First is an understanding of their codes. Very few people know, or for that matter, care that HA0301 is the ISRL classification code for Group DSS. Basically, the code is a key field with some semantic meaning. H = Information Systems, A = Types of Information Systems, 03 = DSS, and 01 = Group DSS. To effectively use these classification codes a knowledge management systems should provide search mechanisms for both the codes and the text representations of the codes. Unfortunately most do not, many only provide search mechanisms for the actual code (i.e. HA0301) and do not provide a search interface to browse the classification code hierarchy.

Second, classification codes, by their very nature in attempting to force a standard set of keywords, are outdated almost immediately. For example, there is no classification codes in the ISRL that represent E-Commerce or XML, but clearly there have been articles written on these topics. As such they must be updated periodically. For example the ACM Classification codes have been updated in 1964 (original), 1982, 1983, 1987, 1991, and 1998 (current). ISRL was created in 1988, and updated in 1993. Recently, (Weber, 2003) Rivard, Barki and Talbot asked *MIS Quarterly* if they would like a revision to the ISRL classification scheme. Interestingly enough the request was denied by the editorial board, and in fact, the ISRL classification scheme has been dropped by *MIS Quarterly*. Some of the factors that led to their decision included: the advances in full-text search engines, database providers providing their own classifications, and the issue of a fast changing IS discipline that requires more frequent updates to its vocabulary.

To overcome the limitations commonly found with classification schemes and address the ambiguity issues with pure free form keywords Chapter Four of this dissertation develops and presents a dynamic, real-time, visual tree-view hierarchy classification scheme architecture that can be utilized in knowledge management systems to improve search efficiencies. This research extends the discussion to include a demonstration that classification schemes can overcome some of these limitations. It also suggests an infrastructure framework that may be more conducive to information retrieval from knowledge management systems.

Recall versus Recognition

As mentioned previously, theory from the psychology reference discipline, specifically the cognitive psychology domain, plays a very important role in the development of this experiment and the formulation of the hypotheses. Recall versus recognition can best be illustrated via typical test questions. Driscoll (2000) illustrates this with the following two questions:

1. What does the word *esoteric* mean?
2. Which of the following words is the best synonym for *esoteric*?
    a. Essential
    b. Mystical
    c. Terrific
    d. Evident

Recall is illustrated by the first question in which few clues are provided. The answerer must formulate an answer, or number of answers and then choose the most plausible. With recognition, the answers are already generated, the answerer need only to recognize which one is correct.

Recall, or rather free recall, has been studied in great depth by psychologists since the late 1800s. This body of research was popularized by Ebbinghaus's 1885 monograph on an experimental investigation into memory (Ebbinghaus, 1913) where recall of nonsense syllables was measured over time. This led to the Ebbinghaus Forgetting Function, which has been reconfirmed time and time again over the decades (Anderson, 1995) and now is considered in psychology as the power law of forgetting. It wasn't until the late 1960s that experiments of recall versus recognition became popular (Bower et al., 1969; Kintsch, 1968; Brown, 1976; Lockart et al. 1976; Anderson et al., 1998; Clark, 1999). The basis of these findings were that in experiments of free recall versus recognition it has been largely shown that recognition routinely outperforms recall in retrieval accuracy, especially when the information was organized hierarchically. In further support of hierarchically-based recognition search strategies Simon (2001) states:

> Thus search processes may be viewed – as they have been in most discussions of problem solving – as processes for seeking a problem solution. But they can be viewed more generally as processes for gathering information about problem structures that will ultimately be valuable in discovering a problem solution. The later viewpoint is more general than the former in a significant sense, in that it suggests that information obtained along any particular branch of a search tree may be used in many contexts besides the one in which it was generated. Only a few problem-solving programs exist today that can be regarded as moving even a modest distance from the earlier more limited viewpoint to the newer one. (p. 127)

Simon (2001) continues developing this discussion by suggesting that the shape of a search design should be hierarchical.

Recall and recognition research has progressed over time from Ebbinghaus' forgetting function through modern theories such as general recognition theory (GRT) and adaptive character of thought theory proposed and advanced by Anderson and colleagues (1997). To help illustrate the historical development of the research leading up to modern day understanding of recall and recognition Table 5 presents a historical timeline. These theories will form the basis for the hypotheses addressing accuracy in Chapter 3.

Table 5

Historical Timeline of Recall versus Recognition Research

Year – Author	Contribution
1885 – Ebbinghaus	Experimental investigation of human memory – recall and the Ebbinghaus forgetting function
1956 – Miller	"The magical number 7 ± 2" and the idea of "chunking"
1962 – Simon	Linking memory searches to hierarchies
1968 – Atkinson & Shiffrin	Atkinson-Shiffron model of human memory (recall processed serially, recognition processed in parallel)
1969 – Bower et al.	Recall versus recognition experiments
1973 – Bower & Anderson	Human Associative Memory (HAM) theory
1986 – Ashby & Townsend	General recognition theory (GRT)
1996 – Anderson	Adaptive Character of Thought theory(ACT)
1997 – Anderson et al.	ACT-R

Effort versus Accuracy

There has been substantial research within information systems on the work effort and accuracy (Todd & Benbasat, 1992; Vessey, 1994; Speier & Morris, 2003). Effort versus accuracy plays an important factor in this research due to the fact that even if a more accurate way of knowledge retrieval could be demonstrated, if it takes substantially more effort to utilize then it may not be effective to implement. In the Todd and Benbasat (1992) article, they discuss effort and decision quality. By introducing decision aides, they sought to determine if there was a difference in effort for those that used decision aids and those that did not.

Speier and Morris (2003) test a research model that examines query interface design for decision performance. Similar to the model tested in this research, the Speier and Morris model provides two interfaces to complete a task. One interface is visual while the other is text based. Their experimental task is searching for a house for possible rental or purchase, whereas the research presented in this dissertation presents several document retrieval scenarios. Their visual interface provides a geographical information systems (GIS) interface for searching for homes that match the stated criteria. Their text-based interface is a traditional query interface found in a typical database application. In their experiment design, Speier and Morris also employ both a low and high level of task complexity. The task complexity was basically manipulated via additional criteria including both "necessary" and "desirable" factors. The top solutions (homes) were determined *a priori*, and subjects were judged against those solutions. Besides capturing the data from the task, they also collected data on subject's spatial ability and work effort. The model in this research manipulates experience and result set size where Speier and Morris manipulate task complexity and spatial ability. Their dependent variable, what Speier and Morris term decision outcomes, was measured by three factors – a work effort instrument, accuracy, and time, whereas in the current study retrieval effectiveness is measured by accuracy time, work effort, and satisfaction. Figure 5 presents the Speier and Morris research model.

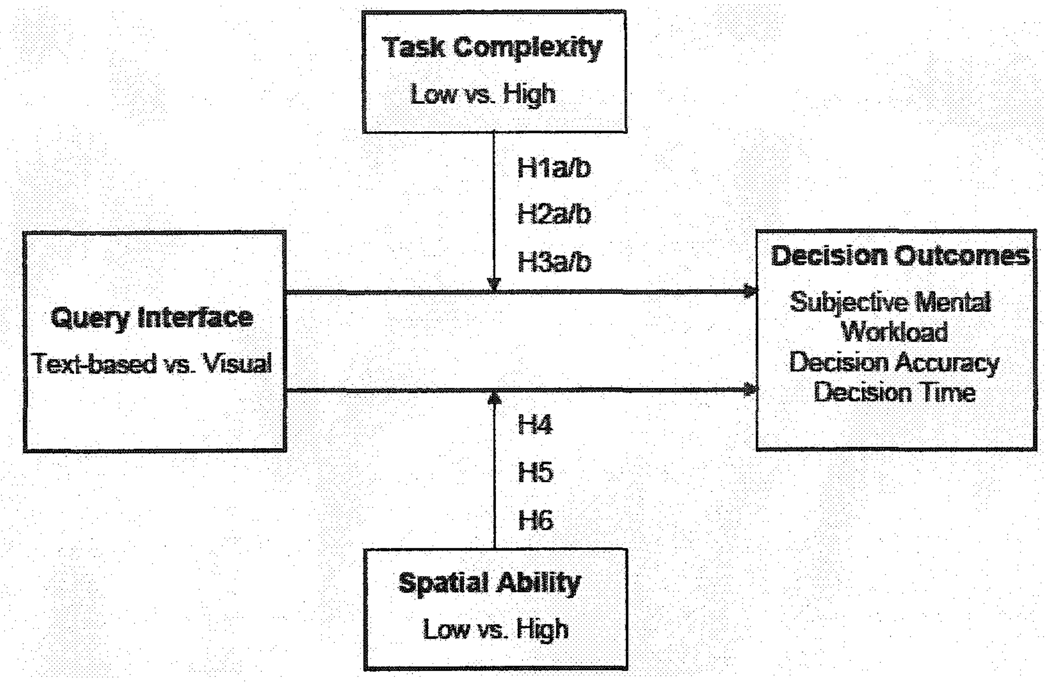


Figure 5. The Speier and Morris query interface research model (Source: Speier & Morris, 2003).

Based on a pool of 372 undergraduate students, with little to know prior database experience, their results found support for:

- Subjective mental workload (work effort) will be lower with visual querying than with text-based querying when task complexity is high.
- Decision accuracy will be higher with text-based querying than with visual querying when task complexity is low.
- Decision accuracy will be higher with visual querying than with text-based querying when task complexity is high.
- Decision accuracy will be higher with high spatial ability decision makers than with low spatial decision makers when using visual querying.

Interestingly, they did not find the predicted results on the following two hypotheses:

- Decision time will be faster with text-based querying than with visual querying when task complexity is low.
- Decision time will be faster with visual querying than with text-based querying when task complexity is high.

The data showed that it actually took less time using the visual interface for low complexity tasks and more time using the visual interface for high complexity tasks. The finding from Speier and Morris (2003) play an important role in the development of the model and hypotheses presented in Chapter 3.

Experience in Information Systems Research

When dealing with cognitive loading issues on task performance it is important to take into account the difference between experienced and novice users. Benbasat and colleagues have studied the theoretical foundations for including experience in IS research models (Dhaliwal and Benbasat, 1996; Gregor & Benbasat, 1999; Mao & Benbasat, 2000). In this research Benbasat and colleagues find support in prior theory that experience can play both a moderating role and have direct effects on information systems research dependent variables. Markus (2001) further differentiates between expert and novices users in developing her theory of knowledge reuse.

Gregor & Benbasat (1999) as well as Bedard (1989) note that finding a generally accepted definition of experience is not a simple task. Even more difficult, however, is finding ways of operationalizing the construct. Some researchers have used professional qualifications and years of experience, while others have used some form of pretest activity to the actual experimental task to be performed. In this research self-reported survey items are used in an attempt to capture user experience. Those items are detailed in Chapter 4.

This chapter has provided a foundational understanding of the breadth of knowledge management research an presents specifically where this research fits within that scheme. Furthermore, it has presented relevant research in support of the differences between data and knowledge retrieval. Cognitive psychology literature has been presented on recall and recognition, effort versus accuracy, and the role of experience in information systems research. The next chapter will develop the research model and based on this survey of prior research and present directional hypotheses that are tested in the experiment conducted for this dissertation.

CHAPTER 3

MODEL AND HYPOTHESES

This section develops the research model used in this experiment based on insights gathered from the review of prior literature presented in Chapter 2. Following the research model, this section then presents the research hypotheses that will be tested by this experiment. The chapter concludes with a summary of all the hypotheses.

Research Model

The recent study by Speier and Morris (2003) provides support for this research in a number ways. First, it suggests that information retrieval is still a topic worthy of further investigation. Visual versus text-based decision support research in information systems is still sparse, with some notable exceptions such as Todd and Benbasat's (1992) decision aids experiment and Vessey's (1991) table versus chart literature review. Second, it supports a number of the experimental design decisions made for this experiment, including a visual versus text-based interface, a manipulation of a moderating experimental variable (they used task complexity, this experiment employs result set size), and the decision to capture work effort data from the subjects. Figure 6 depicts the research model for this experiment.

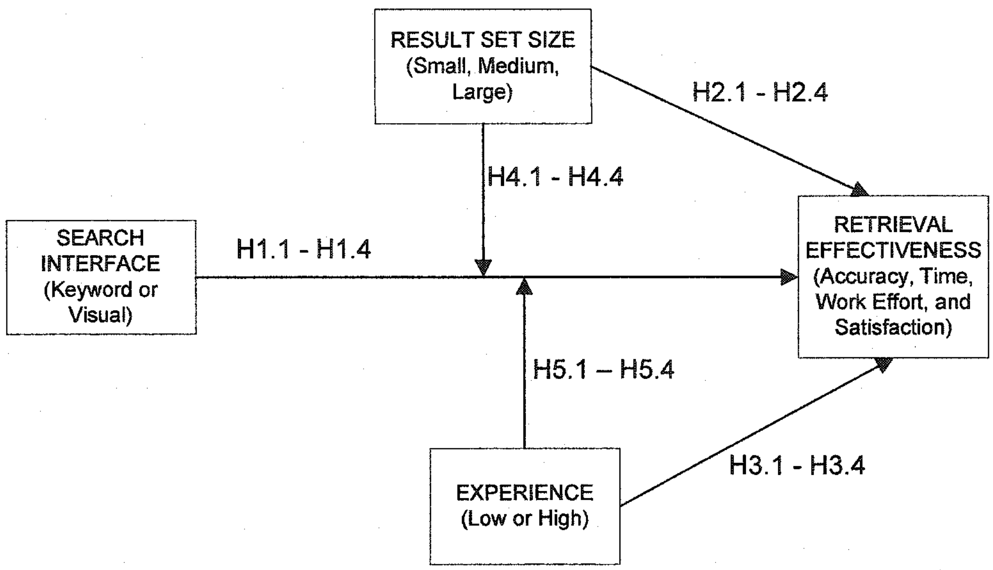


Figure 6. Research model for retrieval effectiveness.

The objective of this research is to focus on the search and retrieval aspects of a typical knowledge management systems. Within that defined scope the following research question is being investigated:

Does the cognitive loading of search mechanisms impact the effectiveness of knowledge retrieval?

Where cognitive loading refers to a recall or recognition paradigm; the search mechanism is either a text-based keyword search function or a visual hierarchical-based search function; and effectiveness is measured via the following four factors: Accuracy, timeliness, work effort, and satisfaction.

Hypotheses

Given the research model based on the research question the following hypotheses are proposed to investigate whether knowledge management systems are employing the most effective means for accurate information retrieval. The hypotheses are broke down as follows:

- A set of main effect hypotheses between the independent variable search interface and the four dimensions of dependent variable retrieval effectiveness (Hypotheses 1.1 – 1.4).
- A set of main effect hypotheses between the independent variable result set size and the four dimensions of dependent variable retrieval effectiveness (Hypotheses 2.1 – 2.4).
- A set of main effect hypotheses between the independent variable experience and the four dimensions of dependent variable retrieval effectiveness (Hypotheses 3.1 – 3.4).
- A set of interaction effect hypotheses between the independent variables search interface and result set size with the four dimensions of dependent variable retrieval effectiveness (Hypotheses 4.1 – 4.4).
- A set of interaction effect hypotheses between the independent variables search interface and experience with the four dimensions of dependent variable retrieval effectiveness (Hypotheses 5.1 – 5.4).

Search Interface on Retrieval Effectiveness.

The first set of hypotheses tests the main effect of search interface on retrieval effectiveness. Directional hypotheses between search interface and each of the four dimensions (accuracy, time, work effort, and satisfaction) of retrieval effectiveness are presented.

General Recognition Theory (Ashby & Townsend, 1986) forms the support for Hypothesis 1.1. Cognitive psychology has extensively studied information retention and retrieval. In particular the debate of recall (keyword) versus recognition (visual) has long been studied. Results have found that recognition is typically a more accurate retrieval method then recall (Bower et al. 1969, Anderson, 1995). In fact, Anderson concludes that, "retrieval of information is facilitated if it is organized hierarchically" (Anderson, 1995 p.223). Hence, the formation of hypothesis 1.1:

Hypothesis 1.1: Knowledge management systems that employ a visual tree-view hierarchy search interface will produce more accurate results than knowledge management systems that employ a text-based keyword search interface.

Or more formally:

Hypothesis 1.1a: Visual tree-view retrieval interface-based knowledge management systems will produce fewer Type I (false-positives) errors than text-based keyword retrieval interface-based knowledge management systems.

and

Hypothesis 1.1b: Visual tree-tree retrieval interface-based knowledge management systems will produce fewer Type II (false-negatives) errors than text-based keyword retrieval interface-based knowledge management systems.

The logic behind hypothesis two can be described via the following scenario. Suppose you are an IS researcher and have a document management system that contains various research articles. You are doing research on the Technology Acceptance Model (TAM) and you would like to find all the articles that have tested TAM (Davis, 1989). Presume that the system actually has 10 articles that truly tested TAM. Through your querying of the system you discover 9 articles. However, upon closer inspection of those articles only seven of them truly tested TAM. You have found only 70% of all the relevant articles. Furthermore, you have produced 2 false positives, which waste your time by not providing the information you requested, and 3 false negatives, or 3 articles that you should have found, but did not find.

A knowledge worker, given enough time, may be able to find all 10 of the relevant TAM articles by running multiple queries. Hypothesis 1.2 suggests that using a visual tree-view hierarchical search interface will typically produce slower, but perhaps more accurate, results based on the knowledge worker browsing through various branch and leaf nodded in their search to discover accurate results. Users will spend less time using a text-based keyword search interface because after exhausting the list of keywords they recall, the knowledge worker will end their search process.

- *Hypothesis 1.2: Users of knowledge management systems that employ visual tree-view retrieval interfaces will spend more time searching than users of knowledge management systems that employ text-based keyword retrieval interfaces.*

Mental work load plays a factor in the day-to-day routines of knowledge workers, to the point where mental fatigue can cause a decrease in performance. Given that assumption it would be wise to develop knowledge management systems that tax the knowledge workers' mental work effort less. Instruments have been developed to test the subjective mental workload of subjects (Hart & Staveland, 1988). These instruments have been applied to the study of subjects involved with using computer based information systems (Speier & Morris, 2003; Morris et al., 1999). Based on the visual aspect of the system, and the workloads associated with recall versus recognition the following hypothesis is posited:

Hypothesis 1.3: Knowledge management systems that employ visual tree-view hierarchy search interfaces will use less work effort than knowledge management systems that employ text-based keyword search interfaces.

Satisfaction has been a popular construct of study within information systems (Parikh et al., 2001; Gelderman, 2002; Doll & Torkzadeh, 1988)
The research conducted in this experiment focuses on two specific aspects of user satisfaction: the process and the results. Primarily the perspective this research takes is based on computer-user feedback. Because of the visual nature of the tree-view hierarchy and the immediate display of results during the search investigation time it is

posited that satisfaction will be higher in the visual tree-view hierarchy based search interface. Hence, hypothesis 1.4 states:

Hypothesis 1.4: Knowledge management systems that employ visual tree-view hierarchy search interfaces will have a higher degree of satisfaction than knowledge management systems that employ text-based keyword search interfaces.

Result set Size on Retrieval Effectiveness.

The second set of hypotheses tests the main effect of result set size on retrieval effectiveness. Directional hypotheses between result set size and each of the four dimensions (accuracy, time, work effort, and satisfaction) of retrieval effectiveness are presented.

Blair (2002a, 2002b) makes a point that document (or knowledge) retrieval systems do not scale as well as data retrieval systems. That is, finding the hire date of the employee with the ID of 123-45-6789 is not any more difficult in an HR database with 1000 employee records or with 100,000 employee records. However, if a user searches a document management system and requests documents on "knowledge management" perhaps 50 out of 1000 (5%) documents are returned and a portion of those are relevant to the knowledge worker. Running that same query on a document management system that has 100,000 documents would return 5000 documents. While a knowledge worker might be able to easily navigate through 50 documents, they will be hard pressed to search through 5000 documents. This problem, Blair suggests, will require the knowledge searcher to submit substantially different, even semantically different, queries than they would on a smaller system.

While this experiment does not provide different sized data sets, it does attempt, in an alternative way, to emulate this issue by manipulating the size of result sets. By experimentally manipulating the result set sizes this experiment emulates large sizes. The amount of data is posited to play a role in the accuracy of information retrieval. Given a small data set it may not matter which retrieval mechanism is used. That is, if there are only 20 documents to sort through in deciding which one may help solve the particular problem at hand, then either architecture (relational or dimensional) could possibly return very similar levels of accuracy. However, as the result set size grows beyond the human capacity to process (Simon, 1976; Baddeley, 1994; Miller, 1956), it is posited that the visual tree-view hierarchy of the dimensional architecture will provide more accurate results over a text-based keyword search of the relational architecture. This is due to the structured nature of the knowledge hierarchies that allow for navigation of the data versus guesswork on the part of the relational system.

Incorporating result set size with the four dimensions of retrieval effectiveness the following hypotheses are presented:

Hypothesis 2.1: As result set size increases retrieval accuracy will decrease.

Hypothesis 2.2: As result set size increases search time will increase.

Hypothesis 2.3: As result set size increases work effort will increase.

Hypothesis 2.4: As result set size increases satisfaction will decrease.

Experience on Retrieval Effectiveness.

Regardless of user interface design of a knowledge management system the underlying technology is always going to be some form of database management system. Given that, it is posited that those with database management system experience will be able to utilize the knowledge management system more effectively, thus resulting in higher levels of information retrieval accuracy. Hence, hypothesis 3.1 states:

Hypothesis 3.1: Subjects with more experience in searching will have higher information retrieval accuracy than subjects with less search experience.

Likewise, with experience come speed efficiencies. The more one has performed a certain type of task successfully in the past the quicker they become at it. Based on the 'practice makes perfect' principle hypothesis 3.2 posits:

Hypothesis 3.2: Subjects with more experience in searching will perform searches faster than subjects with less search experience.

Similar to speed, those with experience will need to exert less effort, thus;

Hypothesis 3.3: Subjects with more experience in searching will use less effort than subjects with less search experience.

Typically those that know how to do something well are generally more satisfied with their work because they lack the frustration factor, thus leading to;

Hypothesis 3.4: Subjects with more experience in searching will be more satisfied than subjects with less search experience.

Hypothesis 3.4 should hold true to a point. For it has been shown in database studies with experienced database developers they would prefer to use SQL language versus a query-by-example user interface. If this holds true in this study then there may be a confounding effect affecting the analysis of this hypothesis.

Search Interface and Result set Size on Retrieval Effectiveness.

The fourth set of hypotheses posits that there is a moderating role played by result set size on the effect that search interface has on retrieval effectiveness. Once again directional hypotheses between search interface and result set size on each of the four dimensions (accuracy, time, work effort, and satisfaction) of retrieval effectiveness are presented.

When it comes to information retrieval size does matter (Blair 2002a, 2002b). Information overload can quickly come into play when performing information retrieval tasks. The experiment tests for this moderating effect of search result set size on the two different search interfaces.

Hypothesis 4.1: As the result set size increases the difference in accuracy between search interfaces will increase negatively.

Hypothesis 4.1 suggests that in small result sets there may be little to no difference in accuracy between search interfaces, where as in large result sets those differences in accuracy will be much greater. As an example, if there are only four articles to be found in a knowledge repository, both search interfaces may produce a very high percentage of those articles. Conversely, as the result set size gets very large the difference in accuracy will increase such that the visual tree-view hierarchical search interface will produce more accurate results than the text-based keyword search interface.

Hypothesis 4.2: As the result set size increases the difference in time between search interfaces will increase positively.

Hypothesis 4.2 is based on the assumption that with larger result sets keyword search users will give up the search faster than those searching via browsing the hierarchy. This should lead to a larger of proportion of time to be spent on the hierarchical search with the increase in size.

Hypothesis 4.3: As the result set size increases the difference in work effort between search interfaces will increase positively.

Similar to the timing issues with hypothesis 4.2, assuming that the keyword searcher will give up earlier on the larger result set sizes this lead to a larger of proportion of work effort to be spent on the hierarchical search with the increase in size.

With respects to the influencing effects of result set size on the search interface with respects to satisfaction it is posited that the difference will increase. This is due to the belief that satisfaction with the keyword search interface will remain relatively constant, whereas the satisfaction of the hierarchy search interface will drop, due to information overload, with large result sets.

Hypothesis 4.4: As the result set size increases the difference in satisfaction between search interfaces will increase negatively.

Search Interface and Experience on Retrieval Effectiveness.

The fifth and final set of hypotheses posits that there is a moderating role played by experience on the effect that search interface has on retrieval effectiveness. Once again directional hypotheses between search interface and experience on each of the four dimensions (accuracy, time, work effort, and satisfaction) of retrieval effectiveness are presented.

Very experienced users will perform well regardless of their tools, thus it is posited that there will be less of a difference between high experienced user and their search interface choice and low experience users. The experiment tests for this moderating effect of experience on the two different search interfaces in terms of accuracy with the following hypothesis:

Hypothesis 5.1: As user experience increases the difference in accuracy between search interfaces will decrease positively.

Hypothesis 5.2 is based on the assumption that with higher experience keyword search users will finish the search task faster than those searching via browsing the hierarchy. This should lead to a larger proportion of time spent on the visual search with the increase in experience.

Hypothesis 5.2: As user experience increases the difference in time between search interfaces will increase negatively.

Similar to accuracy, experienced users are going to put forth less work effort. Assuming they know what they are doing and have done a task over and over to become proficient in it. This phenomenon, known as *automaticity* (Anderson, 1995), supports the following hypothesis:

Hypothesis 5.3: As user experience increases the difference in work effort between search interfaces will decrease negatively.

The influencing effects of experience on the search interface with respects to satisfaction it is posited that the difference will decrease negatively.

Hypothesis 5.4: As experience increases the difference in satisfaction between search interfaces will decrease negatively.

With respects to hypothesis 5.4, this negative decrease may actually show a crossover within the data suggesting that the top most experienced users may in fact be more satisfied with the traditional keyword search interface.

Summary of Hypotheses

Table 6 presents a summary of all the hypotheses to be tested in this experiment.

Table 6

Summary of Hypotheses

H1: Search Interface to Retrieval Effectiveness – Main Effect	Keyword (text-based)	Visual (hierarchy)
H1.1: Accuracy	Less	More
H1.2: Time	Faster	Slower
H1.3: Work Effort	More	Less
H1.4: Satisfaction	Lower	Higher
H2. Result set Size to Retrieval Effectiveness – Main Effect	Small Result sets	Large Result sets
H2.1: Accuracy	Higher	Lower
H2.2: Time	Faster	Slower
H2.3: Work Effort	Less	More
H2.4: Satisfaction	Higher	Lower
H3: Experience to Retrieval Effectiveness – Main Effect	Less Experience	More Experience
H3.1: Accuracy	Lower	Higher
H3.2: Time	Slower	Faster
H3.3: Work Effort	More	Less
H3.4: Satisfaction	Lower	Higher
H4: Search Interface and Result set Size to Retrieval Effectiveness – Interaction Effect		
H4.1: Accuracy	As result set size increases the difference in accuracy increases negatively (higher accuracy for hierarchy in larger result sets)	
H4.2: Time	As result set size increases difference in time increases positively (more time for hierarchy in larger result sets)	
H4.3: Work Effort	As result set size increases difference in work effort increases positively (more effort for hierarchy in larger result sets)	
H4.4: Satisfaction	As result set size increases difference in satisfaction increases negatively (lower satisfaction for hierarchy in larger result sets)	
H5: Search Interface and Experience to Retrieval Effectiveness – Interaction Effect		
H5.1: Accuracy	As experience increases the difference in accuracy decreases positively (higher accuracy for hierarchy in larger result sets)	
H5.2: Time	As experience grows the difference in time increases negatively (more time for hierarchy in larger result sets)	
H5.3: Work Effort	As experience increases the difference in work effort decreases negatively (more effort for hierarchy in larger result sets)	
H5.4: Satisfaction	As experience increases the difference in satisfaction decreases negatively (such that a very experienced user may prefer a keyword search interface over a visual search interface)	

CHAPTER 4

RESEARCH METHODOLOGY

This chapter describes the experiment that tests the hypotheses developed in the previous chapter. This chapter focusses specifically on the experimental design, the variables and their measurements, subject selection and their incentives, and how the instrument was built. This chapter concludes with a walkthrough of the experiment.

Experimental Design

A mixed three factor within subject factorial design was used for this experiment. A laboratory experiment with a 2 x 3 x 2 experimental design was developed to test the hypotheses presented in Chapter 3. An experiment was chosen as the most appropriate methodology for this research because of its strength for showing plausible causal inference (Shadish et al., 2002) and providing the highest level of internal validity (Trochim, 2001). This is done by minimizing the influence of independent variables not pertinent to the investigation (Kerlinger & Lee, 2000). The design of this experiment is considered a mixed factor design (Keppel, 1991) because two factors (Search Type and Experience) are manipulated as a between subject measures. Subjects are randomly assigned only one of the search interfaces and are of only one type of experience (low or high), while the other factor (Result Set Size) is a within subject repeated measure, that is all subjects will be randomly given all three result set scenarios.

Search Interface.

The two factors manipulated in this experiment are search interface (keyword versus hierarchy, between subjects) and result set size (large, medium, and small, within subject). For the search interface manipulation two different search interfaces were developed. The first (control) search interface employs a traditional text-based keyword search interface. Subjects randomly assigned to use this interface enter a keyword or keyword phrase into a textbox on a form and the system returns articles related to that keyword from and underlying relational database management system. This text-based keyword search interface also allowed for the use of wildcard searches. This feature was enabled to ensure the search capabilities of this system were as similar as possible to keyword search interfaces available in today's typical knowledge management systems and search engines available on the Internet. Figure 7 presents the text-based keyword search interface. Appendix A provides all of the screens for the keyword search knowledge management system that is utilized in this experiment.

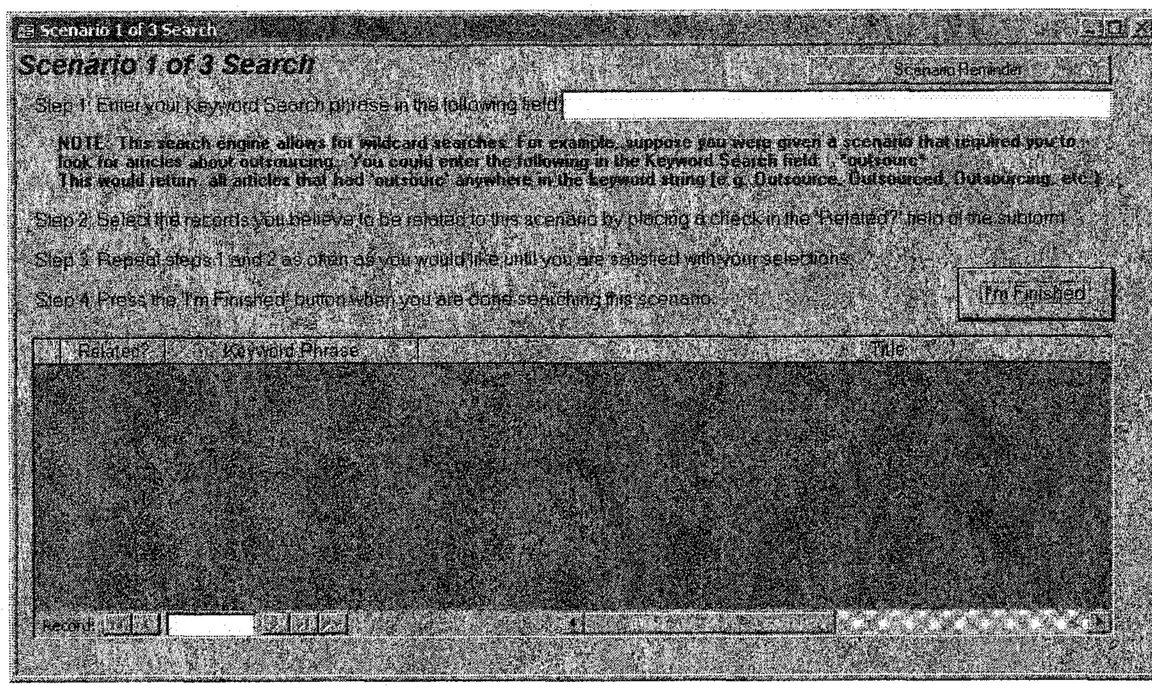


Figure 7. The keyword search interface.

The second (treatment) search interface employs a visual tree-view hierarchy based on a multidimensional data set created from the original relational data set used for the keyword search interface. This visual tree-view hierarchical search allows for knowledge users to navigate up and down a tree-view control to search for related articles. Figure 8 presents the search interface used in the visual tree-view hierarchy treatment group. Appendix B provides all of the screens for the visual search knowledge management system that is utilized in this experiment.

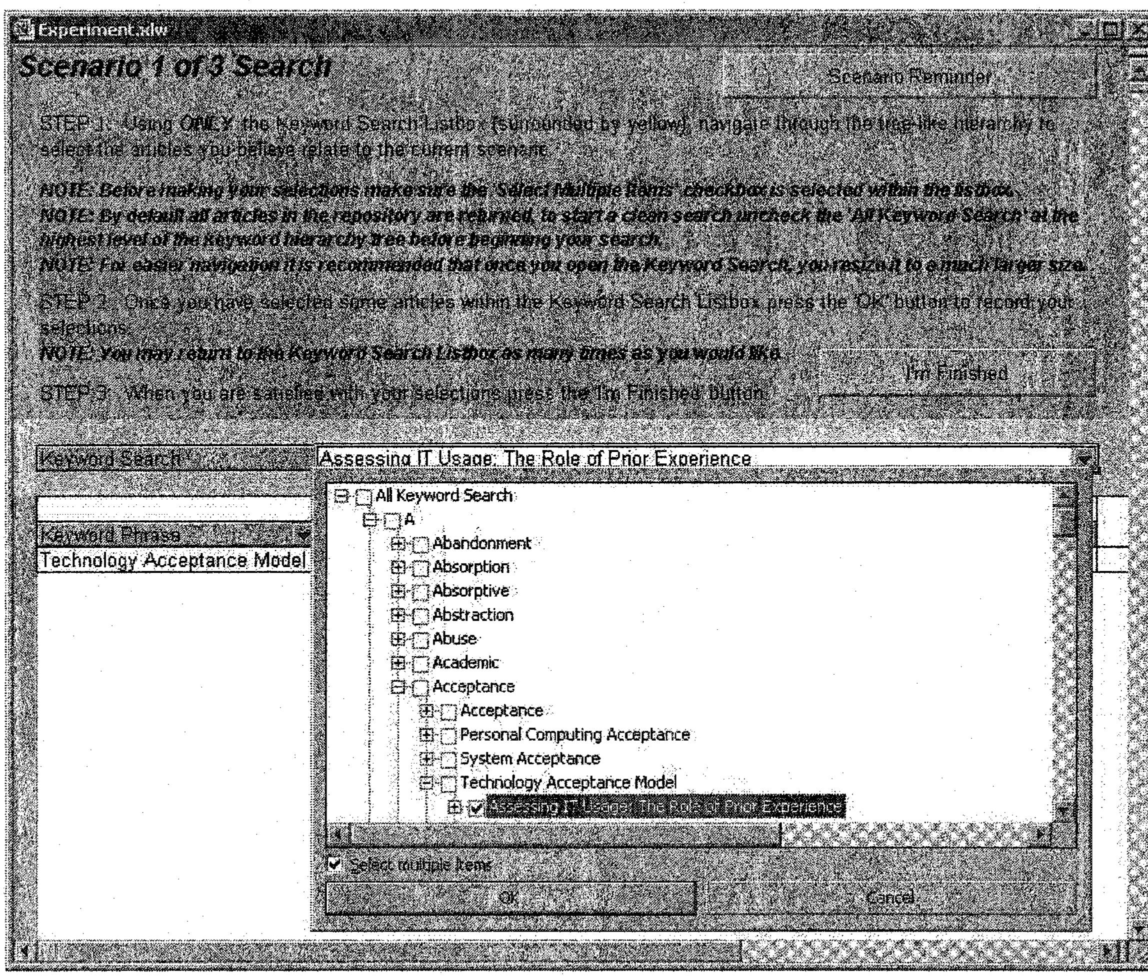


Figure 8. The visual search interface.

Result set Size.

Result set size is the second independent variable in this experimental design. Result set size consisted of three manipulations (large, medium, and small), all of which were measured for each subject in both the control and treatment groups. The result set size variable was balanced via a randomization function within the system so that the scenarios were presented randomly to the subjects. All subjects received all three scenarios, the order of which was randomly generated.

Careful consideration was taken in developing the scenarios so that they would return a wide spectrum of correct results – one large result set, one medium result set, and one small result set. No guidelines were found during the literature review as to what appropriately constitutes a small, medium, or large result sets. For the purpose of this research the correct results for the large scenario constitutes a set which represents 20% of the total population of the original data, or 120 correct journal articles. Correct results for the medium scenario constitutes a set which represents between 6-7% of the total population of the original data, or 40 correct journal articles. Correct results for the small scenario constituted a set which represents between 1-2% of the total population of the original data, or 10 correct journal articles.

Initially nine scenarios (three of each size) were conceived. From those nine, six (two of each size) fully developed scenarios were mocked up, then the research team decided on the final scenarios for each size. The end result was a large scenario based upon the need to discover information on "system design" issues, a medium scenario based upon the need to discover information on "user acceptance" issues, and a small scenario based upon the need to discover information about information systems "risk management" issues. The scenarios are documented later in this chapter under the section entitled *Experimental Walkthrough.*

Experience.

Experience is the third independent variable in this model. Self reported measures were collected in an effort to create an experience measure. Those measures included educational background, search experience, experience with each scenario's topic, and whether or not the subject was considered an IS professional.

Variables and Measurements

The following variables are measured in this experiment in order to test the hypotheses developed in Chapter 3:

- Search Interface (experimentally manipulated independent variable)
- Result set Size (experimentally manipulated independent variable)
- Experience (independent variable)
- Accuracy (research dependant variable)
- Work Effort (research dependent variable)
- Satisfaction (research dependent variable)
- Time (research dependent variable)

Measuring Experience.

Five variables were used to calculate two forms of experience. Two items measured overall information systems experience. Those items are subject self-reported dummy variables (1 = Yes, 0 = No) based on the following two questions:

- *I am considered an IS professional in my workplace?*
- *I hold an IS (or related) degree?*

Three items measured search experience. Those three items utilized 7 point Likert scales for reporting purposes. The questions and their available choices follow:

- *In my workplace I perform searches on large computer information systems (ERP, CRM, Data warehouses, etc.).*
- *In my workplace I perform searches on personal to workgroup size database applications (i.e. Microsoft Access applications, etc.).*

1-Never, 2-Rarely: a couple times a year, 3-Occationally: a couple times a month, 4-Sometimes: approximately once a week, 5-Often: a couple times a week, 6-Daily: approximately once a day, 7-Frequently: several times a day.

- *When I make searches on the Internet (via Google, Yahoo!, MSN Search, etc.) I find what I'm looking for.*

1-Never, 2-Rarely, 3-Occationally, 4-Sometimes: approximately 50% of the time, 5-Most of the time, 6-Almost always, 7-Always)

In addition to the demographic data, the satisfaction, and work effort data that was collected from the subjects, the experiment program also captures the following data:

- Time to perform each scenario search (in seconds);
- Results (number and titles of articles) selected by the subjects for each scenario;
- Searches – in the case of the text-based keyword search interface both the number and actual strings used to search, and in the visual tree-view interface the actual branch/leaf navigation paths are captured.

Measuring Retrieval Accuracy.

In this experiment the dependent variable is information retrieval effectiveness. Effectiveness, in the context of this experiment, is defined as a multidimensional construct that includes the following factors: accuracy, speed, work effort, and satisfaction. For the purposes of this study, information retrieval accuracy is measured by the following method: Two judges with expert knowledge of the data and creators of the search scenarios (the author of this dissertation and his dissertation chair) independently reviewed each article for inclusion or exclusion in the *comparison set.* Following a pseudo-Delphi method approach (Buckley, 1995; Dalkey & Helmer, 1963) articles agreed upon by both judges were included automatically. Next, a list of articles was produced in which one but not both judges agreed should be included in the comparison set. Both judges together worked through this list of articles and came to a consensus on whether the article should be included in the comparison set. This intermediate comparison set was then run against all of the subjects' result sets. A list of articles was produced that included articles that any subject selected but that was not a member of the intermediate comparison set. This list of articles was then evaluated by the judges for any additional articles that should be added to the final comparison set. Upon creating the final comparison set, each subject's result sets was evaluated with the final comparison set to determine a measure of accuracy.

The measure of accuracy not only measures the number of articles correct, but also deducts for incorrect articles. Additionally, the final accuracy score also takes into consideration the number of correct articles that were missed. The formula for calculating accuracy is as follows:

BaseAccuracy = #Correct – (#Incorrect * (TotalPossible / (TotalArticles - TotalPossible))

This BaseAccuracy is then divided by TotalPossible to form an Accuracy percentage score.

Formulating the Accuracy measure this way takes into consideration random chance. If somebody selects *all* the articles his or her score will be zero, likewise if somebody does not select *any* articles his or her score will also be zero. The following examples illustrate the calculation:

The system design (big) scenario has the following characteristics: 589 TotalArticles, 120 TotalPossible articles. If a subject selected all the articles they would have 120 #Correct articles and 469 #Wrong articles.

BaseAccuracy = 120 – (469 * (120 / (589 – 120)))

= 120 – (469 * (120 / 469))

= 120 – (469 * 0.256)

= 120 – 120

= 0

Accuracy = BaseAccuracy / 120

= 0 / 120

= 0%

Similarly, if a subject selected none of the articles they would have 0 #Correct articles and 0 #Wrong articles. This would lead to the following calculation:

BaseAccuracy = 0 – (0 * (120 / (589 – 120)))

= 0 – (0 * (120 / 469))

= 0 – (0 * 0.256)

= 0 – 0

= 0

Accuracy = BaseAccuracy / 120

= 0 / 120

= 0%

Theoretically subjects that produce the perfect results would score the maximum number of points (in the system design scenario there were 120 possible correct articles. The calculations would go as follows:

BaseAccuracy = 120 – (0 * (120 / (589 – 120)))

= 120 – (0 * (120 / 469))

= 120 – (0 * 0.256)

= 120 – 0

= 120

Accuracy = BaseAccuracy / 120

= 120 / 120

= 100%

A subject selecting exactly half the correct articles with no incorrect articles would be rewarded with half the points as a perfect score. This takes into account the Type II, or missing data, error, where the penalty portion of the formula (#Incorrect * (TotalPossible / (TotalArticles-TotalPossible)) takes into account the Type I, or wrong data, error. Calculations for getting exactly half correct are shown below:

BaseAccuracy = 60 – (0 * (120 / (589 – 120)))

= 60 – (0 * (120 / 469))

= 60 – (0 * 0.256)

= 60 – 0

= 60

Accuracy = BaseAccuracy / 120

= 60 / 120

= 50%

This final calculation shows an actual subject score. Subject 114667, for the system design scenario, had 94 #Correct articles and 32 #Wrong articles.

BaseAccuracy = 94 – (32 * (120 / (589 – 120)))

= 94 – (32 * (120 / 469))

= 94 – (32 * 0.256)

= 94 – 8.19

= 85.81

Accuracy = BaseAccuracy / 120

= 85.81 / 120

= 71.51%

MeanAccuracy was then calculated by simply averaging the three accuracy scores together.

Measuring Time.

Time was captured automatically by the experimental system. Time measures began only after the subjects had read the scenario and clicked a the "Start Searching" button which then brought up the Search screen. This method isolated only the time subjects searched for articles and eliminated the varying amount of time subjects needed to read the scenario. Time measurement ended when the subject clicked the "I'm Finished" button on the Search screen and an "OK" button on a confirmation dialog box. If the subjects clicked the "Cancel" button they were able to continue working on the current search task. The confirmation portion was added based on feedback from the pilot study that some subjects accidentally hit the "I'm Finished" button when they were not ready to move on to the next scenario.

Measuring Work Effort.

Work effort is measured using the NASA Task Load Index (TLX) (Hart & Staveland, 1988). This instrument has been used and validated in a number of prior studies (Speier & Morris, 2003; Morris et al., 1999; Grise & Gallupe, 2000; Fisher &

Ford, 1998). The NASA-TLX was chosen because it is quick and easy to administer (Speier & Morris, 2003; Morris et al., 1999; Hart & Staveland, 1988) and human factors researchers (Wierwille & Eggmeier, 1993) show it is appropriate in applied settings, especially within those settings that have low levels of workload.

The NASA-TLX instrument defines work effort as a multidimensional construct consisting of the following six factors: Mental Demand, Physical Demand, Temporal (Time) Demand, Performance Demand, Frustration, and Effort. When completing the NASA-TLX instrument subjects are asked to do a pair-wise comparison on each of the combination of factors. This produces a count of each of the six factors based on the 15 unique comparisons shown in Table 7 below:

Table 7

Pair-wise Comparisons for Measuring Work Effort

Option A	Option B
Effort	Performance
Time Demand	Effort
Performance	Frustration
Physical Demand	Performance
Time Demand	Frustration
Physical Demand	Frustration
Physical Demand	Time Demand
Time Demand	Mental Demand
Frustration	Effort
Performance	Time Demand
Mental Demand	Physical Demand
Frustration	Mental Demand
Performance	Mental Demand
Mental Demand	Effort
Effort	Physical Demand

Next subjects are asked to weight the importance of each of the six factors to the task they just completed. This weighting provides a more sensitive measure to the overall work effort by individualizing the total amount of effort put in by a subject. That is, it provides a weight to each of the factors. The authors of this instrument suggest this weighting be on as large of a scale as possible (Hart & Staveland, 1988). For this experiment a 21-point scale was utilized to allow for scores ranging from 0 to 100 in 5-point increments. Table 8 provides the definitions provided to the subjects for each of the factors.

Table 8

Definitions for the Work Effort Factors

Factor	Definition
Mental Demand	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical Demand	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Time Demand	How much time pressure did you feel due to the rate or pace at which the task occurred? Was the pace slow and leisurely or rapid and frantic?
Performance	How successful do you think you were in accomplishing the goals of the task? How satisfied were you with your performance accomplishing these goals
Effort	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration	How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

With these two sets of data the following calculations are performed to arrive at individual scores for each of the factors and a weighted overall score for work effort.

FactorScore = CountfromCompairson * FactorWeight*ScaleIncrement

OverallWorkEffort = (SumofFactorScores) / 15

As a numerical example, subject 502819 was randomly assigned the keyword search interface. For the user acceptance (medium) search scenario the following data was collected:

MentalCount = 3	MentalWeight = 15
PhysicalCount = 0	PhysicalWeight = 0
TimeCount = 2	TimeWeight = 5
PerformanceCount = 5	PerformanceWeight = 17
FrustrationCount = 1	FrustrationWeight = 3
EffortCount = 3	EffortWeight = 14

Yielding the following individual factor scores:

MentalScore = 3 * 15 * 5 = 225

PhysicalScore = 0 * 0 * 5 = 0

TimeScore = 2 * 5 * 5 = 50

PerformanceScore = 5 * 17 * 5 = 425

FrustrationScore = 1 * 3 * 5 = 15

EffortScore = 3 * 14 * 5 = 280

With each of the individual factor scores it is now possible to calculate an overall work effort score for the medium search task for subject 502819.

OverallWorkEffort $= (225 + 0 + 50 + 425 + 15 + 280) / 15$

$= 995 / 15$

$= 66.33$

This score is only for the work effort for one (the user acceptance scenario) search task. This instrument was administered directly after the completion of each search task for a total of three times per subject. This was done to allow for comparisons between each of the different search tasks.

Measuring Satisfaction.

User satisfaction is a perception variable that many different researchers have defined and measured in a number of different ways (Parikh et al., 2001). To evaluate overall user satisfaction with each of the search interface mechanisms the End-User Computing Satisfaction (EUCS) instrument was used. The EUCS survey instrument was developed and validated by Doll and Torkzadeh (1988; Doll et al., 1994). This instrument was chosen because it measures the satisfaction with an individual application and not with the information in general.

As with work effort, satisfaction is defined as a multidimensional construct. The five factors that define satisfaction are: Content, Accuracy, Ease of Use, Format, and Timeliness. Twelve questions were asked, 4 for Content, and 2 each for the remaining factors, using a 5-point Likert scale, ranging from 1 = non-existent to 5 = excellent.

Individual factor scores were calculated by averaging the scores for each of the questions for the individual factor. Then an overall satisfaction score was calculated by summing each of the individual factor scores for a possible range of 5 to 25.

Like the utilization of the NASA-TLX instrument, the EUCS instrument was administered directly after each search task for a total of three times for each subject. In this way computing satisfaction can be measured between subjects as well as across result set size.

Subject Selection

Students completing various masters' degrees in business (Master's of Business Administration or Master's in Accounting Information Systems) at a major university in the southwestern United States are the primary subject group of this experiment. This subject pool was selected as they are most likely to exhibit traits of the common knowledge worker. Many of them have previous or current knowledge work experience. Most have completed a course in database management either at the undergraduate or graduate level while others have not. Furthermore, because they were Business School students they were believed to have some familiarity of the scenarios presented in the experiment as well as some practice with searching for business information systems literature. To increase the number of subjects a group of senior business undergraduates from the same university was also utilized in this experiment.

Incentives and Motivation.

Incentives are made available in two forms, performance-based (Marsden, et al., 2002; Smith, 1976) and non-performance-based (Jordan, 1986). As is common practice in much of the behavior research field, course credit was given to those who participated in the experiment (non-performance-based). In the one case in which the experiment was held for an MBA student group, rather than in a class, a random prize drawing (non-performance-based) was held where the recipient was awarded a keychain USB JumpDrive. Additionally, all participants were given the opportunity to participate in a performance-based incentive. A $100 reward for the highest accuracy rate was announced during the introduction of the experiment. While this reward was not large it was meant to simulate a typical reward for performance that you might find in a sales or consulting organization. This incentive strategy is similar to that employed by Marsden et al. (2002), and is based on Smith's (1976) induced value theory.

Instrument Development

A common practice for most knowledge management systems is to rely on keywords as the primary mechanism for search and retrieval. While this has suited knowledge workers in the past, it is not without its limitations. To demonstrate these keyword search limitations a database was developed to analyze all of the articles that have appeared in *MIS Quarterly* during its 27 year history. The dataset includes 611 articles with 1770 keywords occurring 2857 times, and 610 classification categories occurring 2280 times. Results can be found in LaBrie and St. Louis (2003), findings from that study along with the research question outlined in Chapter One provided the motivation for this experiment.

To operationalize this experiment two forms of a representative knowledge management system were developed. These knowledge management systems were developed as representative document management systems. Document management systems are often examined in IS research due to their popular use throughout the business and academic world (Sprague, 1995; Lambrix & Shahmehri, 2000; Blair, 2002a, 2002b).

Each system hosted the same data, albeit in different formats. The data set was built around the entire collection of *MIS Quarterly* journal articles stored in electronic (Adobe® Portable Document Format – pdf) format along with some associated meta data. In all, 611 journal articles, ranging from volume 1, issue 1 in March 1977 through volume 27, issue 3 in September 2003 – a 27 year view of IS research, is used as the data set.

Keyword Instrument Development.

The first knowledge management system developed (hereto after referred to as KMS-Keyword), was developed with Visual Basic for Applications (VBA) in Microsoft® Access XP. KMS-Keyword is a relational data storage architecture that is based on an application that houses not only the actual article, but also all of the data and metadata that is typically found in any traditional document management system. This data includes: author, and author information (degree, granting university, current location, etc.) title, abstract, issue information, keywords, classification codes (categories), and type of article (research, editorial, etc.). Figure 9 shows the conceptual schema of the relational database that KMS-Keyword is based on.

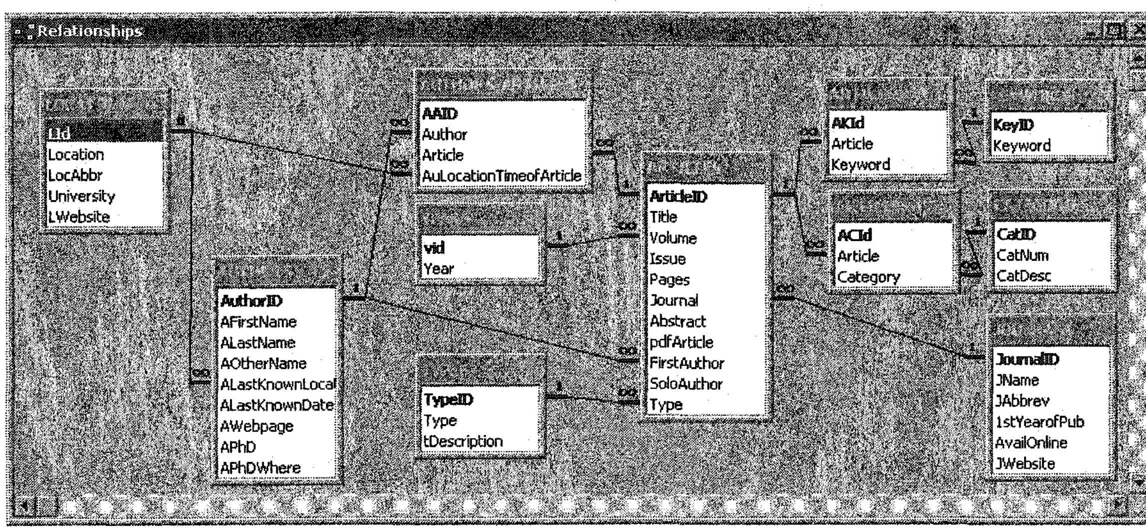


Figure 9. A schema for an IS journal knowledge management system.

For purposes of this experiment, a scaled down version of this document management knowledge management system was used. This decision was made so that the experimental database would be more portable. Primarily, author information and the actual electronic documents (the PDF files) were removed as the experimental user interface did not involve their usage. The final conceptual schema used for the experiment is shown in Figure 10.

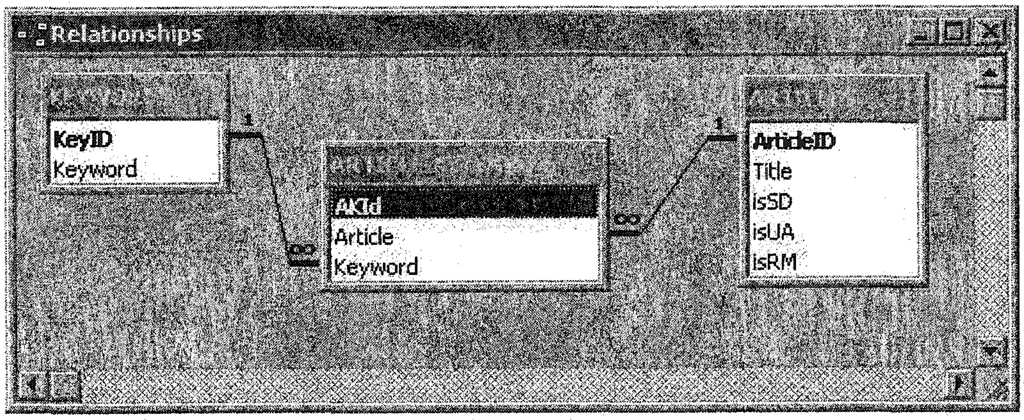


Figure 10. A relational model for the keyword search system.

Basically all that was necessary for the experiment was the keyword data, some article data, and the many-to-many relationship between keywords articles.

Visual Instrument Development.

The second form of the document management knowledge management system (hereto after referred to as KMS-Hierarchy) that was developed for this experiment was based on a multi-dimensional data storage architecture. The same data used in the KMS-Keyword was used in the population of KMS-Hierarchy. Where relational systems have been optimized for data input, multidimensional systems have been optimized for information output (Inmon, 1996; Hoffer et al., 2002). Kimball describes the dimensional model as the only practical way to present data to the end-users (Kimball, 1996). Figure 11 shows the dimensional model for the multidimensional KMS-Hierarchy.

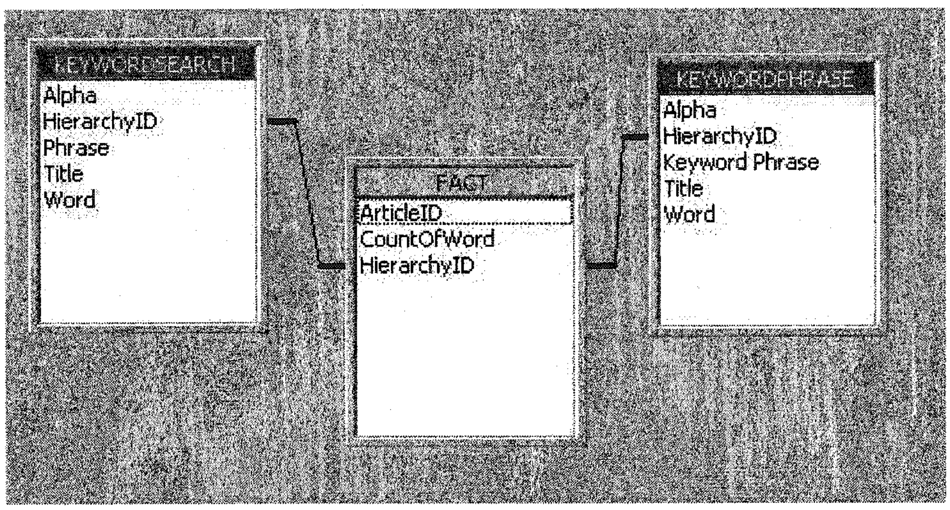


Figure 11. A dimensional model for the visual search system.

While this dimensional model looks very relational in its appearance, the underlying storage mechanism employs a multidimensional online analytical processing (MOLAP) cube. The cube was created within Microsoft SQL Server 2000, Analysis Services. One of the key advantages of using the MOLAP storage engine is its allowances for hierarchies and aggregates. Once the cube was built, subjects are able to browse the data via the following hierarchy: alphabetic letter, keyword, keyword phrase, and title of article. Figure 12 shows a sample of the keyword hierarchy.

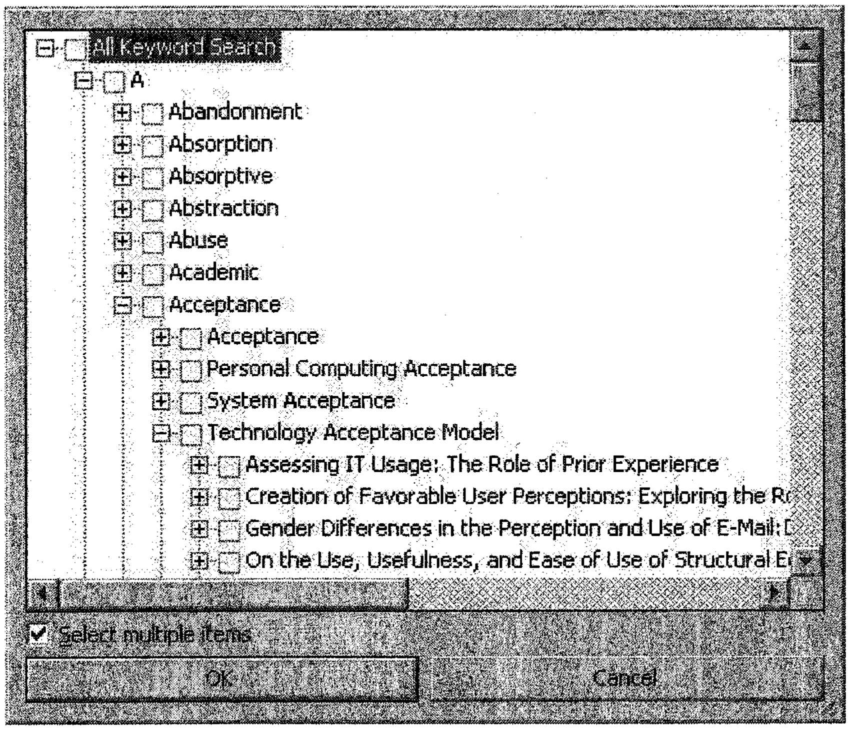


Figure 12. The intelligent keyword hierarchy.

Developing the Keyword Hierarchy.

Various classification schemes were examined for use as a potential hierarchy for this experiment. Both the ACM and ISRL classification schemes posed limitations in that neither where used consistently throughout the entire timeframe of the data set. Additionally, as mentioned in the literature review, the limitation of classification scheme becoming quickly outdated (Weber, 2003) and lacking new terminology (i.e. XML, electronic commerce, and web services to name a few), caused the researcher to develop a new classification scheme. The keyword classification scheme is a dynamic real-time keyword based classification scheme. Keyword phrases are taken from the keywords of articles as they are entered into the system. The keyword phrases are then broken down to their word elements, trivial words are removed such as: a, the, and, etc. The individual words are then linked back to their original keyword phrase and the article they are associated with. The results are a searchable hierarchy that allows a user to search via any of the words within the keyword phrase. For example if somebody is looking for articles on "Technology Acceptance Model" they would find them in three different areas of the hierarchy: 1) Under A, Acceptance, Technology Acceptance Model (as shown in Figure 12 above) or 2) Under T, Technology, Technology Acceptance Model, or 3) Under M, Model, Technology Acceptance Model. The benefit of this strategy allows for knowledge workers to search for keywords, regardless of their position in the keyword phrase. This can be thought of as the visual tree-view hierarchical equivalent of allowing wildcard searches within a text-based keyword search interface.

Experimental Walkthrough

This laboratory experiment will be conducted electronically via personal computers in a windows-based environment. The researcher and any potential facilitators will be provided with a script for conducting the experiment – script to come as an appendix. Each subject will be presented with a computer with the appropriate hardware and software necessary to complete the experiment. Subjects will be randomly assigned to perform searches from one of two programs. The programs were developed to be exact in every way except for the search interface. The procedure to be followed when conducting the experiment follows:

1. After a brief introduction by the experimenter including a welcome, a thank you, and a brief introduction to the experiment the subjects are instructed to click on an icon on the desktop that starts the experiment.
2. The subjects are then presented with a welcome screen that explains the purpose of the experiment. On this screen they are able to voluntarily exit without beginning the data collection, or to continue on with the experiment.
3. The subjects are then presented with a screen that collects demographic information. This data collected on this screen includes: age (year of birth), gender, education, work experience, and search experience. From this screen subjects can return to the introductory screen, or can choose to begin the experiment.
4. By beginning the experiment the subjects are then provided with three scenarios assigned in random order. First a scenario is presented to them, and then the search interface is made available. This constitutes six screens, one for the introduction of each scenario and one for the search interface for each scenario. The search interface screen is the same screen repeated for each scenario. That is, the subjects are given either the text-based keyword search screen for all three of the scenarios or they are given the visual tree-view hierarchical search screen for all three scenarios.
5. After the completion of each scenario the subjects are presented with a screen that measures the work effort for the search task just completed. Following the screen

that measures work effort, the subjects are then presented with a screen to measure their satisfaction with the search task just completed.

6. Upon completing all three scenarios and searches, the subjects are then presented with a closing screen that thanks them for their participation and collects data on their experience with each scenarios topic. Also made available on this screen, is a place for optional comments and the choice receiving the results of the experiment. By choosing the exit button on this screen the experiment program is closed and a data file necessary for analysis of the results is created for the researcher.

The actual computer screens are provided in Appendices A and B. The following sections provide the text of the screens and present the search interfaces.

The Welcome Screen

Welcome and thank you for participating in this experiment examining information retrieval accuracy from knowledge management systems. Your involvement in this study is greatly appreciated.

This experiment is designed to test the effectiveness of an information system interface; not your individual ability. As such, all data collected during this experiment is presented anonymously to the researchers.

The setting of this experiment is that you are a knowledge worker that has access to an information repository that contains a number of journal articles related to your field. You will be presented with a simplified user *interface that will allow you to perform keyword searches for articles that may be related to the scenarios presented to you (much like an online library indexing service). Once you have searched the system and articles have been returned, you will then be able to select the articles you think will be helpful to that scenario.*

During this experiment you will be given three (3) different scenarios in which you must search for articles relating to that specific topic. Because there is a timing component to this experiment, the researchers ask that you complete the tasks without interruption. Please feel free to take as long as you want, just don't perform any other activities (email, chat, Internet browsing, etc.) during the duration of this experiment.

To begin offering demographic data and to continue on to the experiment please click the 'Next>>' button, to exit without participating in the experiment please click the 'Exit' button.

The Demographics Screen

The demographics screen asks for the following information:

- *Please enter the year you were born*: (Choices: open textbox that accepts years from 1920-1988).
- *Gender*: (Choices: Female or Male).

- *Please select the highest level of education you have attained:* (Choices: High School (or equivalent), Bachelor's Degree, Master's Degree, or Doctorate Degree).
- *I am considered an IS Professional in my place of work.* (Choices: Yes/No checkbox).
- *I hold an IS (or related) degree.* (Choices: Yes/No checkbox).
- *In my workplace I perform searches on large computer information systems (ERP, CRM, Data warehouses, etc.).* (Choices 7-point Likert scale: 1-Never, 2-Rarely: a couple times a year, 3-Occationally: a couple times a month, 4-Sometimes: approximately once a week, 5-Often: a couple times a week, 6-Daily: approximately once a day, 7-Frequently: several times a day).
- *In my workplace I perform searches on personal to workgroup size database applications (i.e. Microsoft Access applications, etc.).* (Choices 7-point Likert scale: 1-Never, 2-Rarely: a couple times a year, 3-Occationally: a couple times a month, 4-Sometimes: approximately once a week, 5-Often: a couple times a week, 6-Daily: approximately once a day, 7-Frequently: several times a day).
- *When I make searches on the Internet (via Google, Yahoo!, MSN Search, etc.) I find what I'm looking for.* (Choices 7-point Likert scale: 1-Never, 2-Rarely, 3-Occationally, 4-Sometimes: approximately 50% of the time, 5-Most of the time, 6-Almost always, 7-Always).

The Scenario Screens

The following three scenarios are presented to all participants of the experiment. The scenarios are randomized to control for any learning effect.

System Design Scenario.

In this scenario, suppose you are a manager in a sizable information technology department within a large corporation. Assume your department is responsible for a large number of internal IT application development projects. Many of your projects tend to have problems that you would like to see alleviated. You believe that many of these problems could have been avoided by better design during the information systems development process. Before moving forward with any new projects you would like to learn about ways to more effectively design information systems.

Using the search interface provided on the next screen, seek out articles that will help you in learning about this issue. When you find an article that you believe is related to the topic please place a check in the box provided.

To begin timing of this task press the 'Start Searching' button and the search form will appear. As soon as you are satisfied with your selections press the 'I'm Finished' button.

Good Luck with your searches!

User Acceptance Scenario.

In this scenario, suppose you are working on a project for a self-monitoring healthcare application in which you must implement a new computer system that patients will need to use in their home. You would like to learn more about what causes end-users to accept new information systems.

Using the search interface provided on the next screen, seek out articles that will help you in learning about this issue. When you find an article that you believe is related to the topic please place a check in the box provided.

To begin timing of this task press the 'Start Searching' button and the search form will appear. As soon as you are satisfied with your selections press the 'I'm Finished' button.

Good Luck with your searches!

Risk Management Scenario.

In this scenario, suppose you are a senior manager of an IT organization in a company that is heavily dependent on the use of computing technology. Due to the recent floods of computer viruses, Internet worms, and hackers trying to gain access to customer records *your systems have come under scrutiny by the top management team and board of directors. Because of this, you decide to seek a greater understanding of how to mitigate risk in information systems to better safeguard against these dependencies.*

Using the search interface provided on the next screen, seek out articles that will help you in learning about this issue. When you find an article that you believe is related to the topic please place a check in the box provided.

To begin timing of this task press the 'Start Searching' button and the search form will appear. As soon as you are satisfied with your selections press the 'I'm Finished' button.

Good Luck with your searches!

The Search Screens

There are two different search mechanisms within the design of this experiment. The first is a text-based keyword search interface built on top of a relational database engine. The keyword search interface was presented previously in Figure 7. The second search mechanism is a visual tree-view search interface built on top of a multidimensional database engine – the data source being the same data stored in the relational database engine used in the text-based keyword search interface program. The visual tree-view search interface was presented previously in Figure 8.

The Work Effort Screens

A work effort screen is presented immediately after the completion of each search task. The work effort screen presents the following instructions and collects the following data for calculating work effort.

As a part of this experiment data needs to be collected about the workload you experienced in performing the search task. For each pair, please select the member that was the more important contributor to the workload level of the search task just performed. Next, please select a point on the scale that best represents the magnitude of each factor for the search task you just performed. Definitions of the terms are available by clicking their respective button.

- Pair-wise comparisons of the six factors (mental demand, physical demand, temporal demand, frustration, effort, and satisfaction) that make up the work effort construct.
- 5% increment scale (from 0 to 100) on the user's perception on how important each factor (mental demand, physical demand, temporal demand, frustration, effort, and satisfaction) should be weighted for that specific task.

The Satisfaction Screens

A user satisfaction screen is presented immediately after the completion of each work effort screen. The satisfaction screen presents the following instructions and questions for calculating satisfaction.

As a part of this experiment data needs to be collected about the satisfaction you experienced in performing the previous search task. For each of the following questions please rate the level of satisfaction on a scale from 1 to 5.

- *Does the system provide the precise information you need?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).
- *Does the information content meet your needs?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).
- *Does the system provide reports that seem to be just about exactly what you need?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).
- *Does the system provide sufficient information?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).
- *Is the system accurate?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).
- *Are you satisfied with the accuracy of the system?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).
- *Do you think the output is presented in a useful format?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).

- *Is the information clear?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).
- *Is the system user friendly?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).
- *Is the system easy to use?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).
- *Do you get the information you need in time?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).
- *Does the system provide up-to-date information?* (Choices 5-point Likert scale: 1-non-existent, 2-poor, 3-fair, 4-good, 5-excellent).

The Closing Screen

Thank you for your participation in this experiment. To successfully complete this experiment the following three questions must be answered. After answering these questions you optionally have the opportunity to provide feedback to the researchers and may enter your email address if you would like to be notified of the results from this experiment.

The closing screen then asked the subject for their level of experience with each of the subject matters presented in the search tasks with the following questions:

- *I am familiar with the topic of: Information Systems Design.* (Choices 7-point Likert scale: 1-Strongly disagree, 2- Disagree, 3- Somewhat disagree, 4- Neither agree nor disagree, 5- Somewhat agree, 6- Agree, 7- Strongly agree).
- *I am familiar with the topic of: User Acceptance of Information Systems.* (Choices 7-point Likert scale: 1-Strongly disagree, 2- Disagree, 3- Somewhat disagree, 4- Neither agree nor disagree, 5- Somewhat agree, 6- Agree, 7- Strongly agree).
- *I am familiar with the topic of: Information Systems Risk Management.* (Choices 7-point Likert scale: 1-Strongly disagree, 2- Disagree, 3- Somewhat disagree, 4- Neither agree nor disagree, 5- Somewhat agree, 6- Agree, 7- Strongly agree).
- *OPTIONAL: If you would like to leave any comments about this experiment for the researchers please provide them in the space provided below.* (Choice: Comment box made available).
- *OPTIONAL: I would like to be notified about the results of this experiment.* (Choice: Yes/No checkbox, upon checking 'Yes' a '*Please enter your email address*' field becomes available).

CHAPTER 5

ANALYSIS AND RESULTS

This chapter details the analysis and reports the results of the experiment. Prior to running the full experiment a pilot study was performed. Pilot study results are reported first followed by the analysis and results of the experimental study for each of the dimensions of retrieval effectiveness (accuracy, time, work effort, and satisfaction). Descriptive statistics and GLM equations are presented for each dimension for each scenario, followed by a pooled model and a repeated measures model for each dependent variable factor.

Pilot Study

A pilot study was conducted with 10 information systems doctoral students. The pilot was meant to test the face validity (Kerlinger & Lee, 2000) of the experiment. Based on feedback from the pilot study a number of minor design changes were made to decrease the complexity of the search task and improve the usability of the interface. Based on informal interviews with the pilot study subjects it was concluded that the scenarios presented in the experiment were understandable and that the search interfaces performed as expected.

An additional objective of the pilot study was to confirm the general directions of the hypotheses of this study (i.e. the visual tree-view hierarchy search interface would produce more accurate results, take longer time, and be more satisfying). While statistical significance could not be concluded with these few subjects, each of the general trends expected were confirmed by the pilot study. Table 9 presents the results based on the data collected from the pilot study.

Table 9

Pilot Study Results

Measure	Visual	Keyword
Big Count	100.50 (22.33)	4.50
Med Count	13.25	3.17
Sml Count	6.75	1.33
Big Time	209.25	264.67
Med Time	405.75	283.83
Sml Time	210	239
Total Time	*825 (13m45s)*	*787.5 (13m7.5s)*
Process-Satisfaction	5.42	2.56
Results-Satisfaction	5.50	2.78
Overall-Satisfaction	5.46	2.67

For the evaluation of these results a proxy of a (mean) count represents accuracy. This is because the rating of the articles by the judges had yet to be completed. In all scenarios the results show that the subjects who used the visual tree-view hierarchy search interface found more articles. Furthermore, the satisfaction of the visual tree-view hierarchy search interface system (both process and results) was considerably higher. On a cautionary note the time spent between the two systems did not vary much. The cause of this may be the difficulty in the first implementation of the keyword search interface. Changes made to the search interface portion of the program subsequent to the administration of the pilot study simplified the keyword search user interface. In the final experiment subjects simply typed in a keyword or keyword phrase, including wildcard searches, and hit enter to have results returned. During the pilot study subjects had to press a button to go into a special search mode and then press another button in order to return the results. Pilot study subjects that used the keyword search interface found that process awkward so it was changed to better represent a more typical keyword search interface. The NASA-TLX work effort instrument was not implemented at the time of the pilot study, so there are no results for work effort to report.

Experimental Data Collection

Data for this experiment was collect during four different sessions. Each session was proctored by the researcher. In each session the following procedure was followed:

1. A general introduction set the context of the experiment.
2. A demonstration of both search interfaces was presented.
3. The experiment was run.
4. The subjects were debriefed on the experimental measures.

For the general introduction a 20 minute presentation introduced the subjects to issues surrounding the impending Sarbanes-Oxley Act. This topic was selected as timely and an example of electronic document management issues faced by today's knowledge management systems. Appendix C provides the slide deck utilized for the introduction of the experiment.

Upon setting the stage with the general introduction, the researcher then proceeded with a demonstration of both systems. The keyword search system was presented first, with the experimenter walking through each screen. The actual scenarios were not shown. Instead a scenario based on a database literature search was presented. After the keyword interface was presented, the researcher presented the visual search interface, using the same database literature scenario. Subjects were then reminded that this was a timed, anonymous experiment, but if they chose to participate in the $100 cash prize for highest accuracy that they would need to provide their email address on the closing screen of the experiment. Subjects were given as much time as they wanted to complete the experiment. The range of time taken to complete the experiment was 20 minutes to approximately one hour. Those that took the most time took approximately one hour.

Prior to the commencement of the experiment the subjects were asked to return at a specific time to be debriefed. The debriefing was conducted in the same manner for each of the sessions for master students. The undergraduate students received an abbreviated version of the debriefing upon the award of the prize. The debriefing slide deck is available in Appendix D. The debriefing described the research model and presented the hypotheses at a high level. The debriefing took approximately 20 minutes. The entire study took approximately two hours.

Demographic Data.

Table 10 presents the descriptive statistics for the demographic data collected in this experiment.

Table 10

Demographic Data

Variable	Response Category	Number of Subjects	Accuracy Mean	Time Mean	Effort Mean	Satisfaction Mean
Gender						
	Female	26	23.85%	984	59.85	15.80
	Male	49	23.57%	1004	58.79	14.92
Age						
	<21	0	---	---	---	---
	21-25	29	21.29%	753	56.40	14.99
	26-30	30	27.26%	1207	62.99	15.28
	31-35	11	21.86%	1049	57.58	14.35
	36+	4	22.78%	1196	59.24	16.73
Education						
	Undergrads	33	20.12%	754	57.13	15.10
	Graduate	42	26.45%	1188	60.58	15.21
Anonymity						
	No	59	24.55%	981	58.19	15.44
	Yes	16	20.40%	1056	63.00	14.40

Independent sample *t*-tests were conducted to test for significant differences between the samples of the two retrieval methods for gender, age, education, and anonymity. No significance was found for gender, age, or anonymity.

Handling of Outliers.

A total of 79 subjects participated in the experiment over the course of four sessions. Four subjects (298825, 433651, 752300, and 893368) were dropped from the analysis because each chose to skip at least one of the scenarios, leaving 75 usable cases for analysis. Box-plots were utilized to identify other potential outliers. The Box-plots only identified subjects with extremely high accuracy rates. Because those were actual scores it was decided to leave those subjects in the data set.

Replacing Experience with Education

Upon analysis of the data collected for experience it was found that there were little to no statistically significant variables that could act as an appropriate measure for experience. Analysis was performed on the following self reported scale items to test for significance:

- I have an IS degree.
- I am considered an IS professional at my place of work.
- When I search large IS systems I find what I am looking for.
- When I search small IS systems I find what I am looking for.
- When I search the Internet I find what I am looking for.
- I am familiar with the topic {System Design, User Acceptance, Risk Management}.

Only *IS professional* showed significance, and it was heavily influenced by the control variable for education. Further analysis found only one undergraduate considered her/himself to be an IS professional; all other self-reported IS professionals were graduate students. Interestingly, there were very significant differences between the undergraduate and graduate population. So much so that it was decided to modify the model and substitute education for experience. This makes some sense as the graduate students tended to be older – suggesting the possibility for more work experience, were much more likely to have IS degrees, and had a much higher level of reporting themselves as IS professionals.

Factor Analysis of Experience Items.

Further statistical analysis was performed to investigate whether or not the experience items were actually loading on a single experience factor. Using SPSS 12.0 a Principal Components Analysis with a Varimax with Kaiser Normalization rotation method showed three experience factors were actually measured. The rotations converged after four iterations. Table 11 presents the results of the factor analysis.

Table 11

Factor Analysis of Experience

	Component		
	1	2	3
SDExpert	0.144	0.873	-0.04
UAExpert	-0.021	0.883	0.105
RiExpert	-0.048	0.836	0.089
Graduate	0.861	-0.087	-0.243
Educ	0.792	-0.124	-0.168
ISProf	0.738	0.143	0.226
ISDeg	0.733	0.165	0.256
UseISBig	0.011	-0.113	0.862
UseISSml	0.012	0.235	0.746

The three experience factors that can be derived from this factor analysis can best be described as:

- Content experience (loaded on by SDExpert, UAExpert, RMExpert).
- Educational experience (loaded on by ISDeg, ISProf, Educ, Graduate).
- Search experience (loaded on by SearchBig, SearchSml).

Search experience and content experience had now explanatory power. For educational experience only the *graduate* factor had any explanatory power. Hence, the graduate educational experience factor is used as a proxy for experience. Based on these finding, all analysis from this point forward is done with education as a proxy for experience. The final model presented in the *Discussion* section of Chapter 6 reflects this change. Further discussion of this change is also presented in the *Limitations of the Study* section in Chapter 6.

Analysis of Accuracy

Table 12 provides descriptive statistics on the dependent variable accuracy; including accuracy based on result set size, retrieval method, and education.

Table 12

Descriptive Statistics for Accuracy

	Overall	(n=75)		Visual	(n=40)		Keyword	(n=35)
Accuracy	Mean	Std Dev		Mean	Std Dev		Mean	Std Dev
Sml	38.45%	0.2106		45.38%	0.2008		30.53%	0.1954
Med	17.04%	0.1833		21.56%	0.1899		11.87%	0.1631
Big	15.57%	0.1635		19.92%	0.1872		10.60%	0.1150
Total	23.69%	0.2135		28.96%	0.2240		17.67%	0.1840
				Graduate	(n=42)		Undergrad	(n=33)
			Accuracy	Mean	Std Dev		Mean	Std Dev
			Sml	39.95%	0.2020		36.54%	0.2229
			Med	19.95%	0.1836		13.34%	0.1788
			Big	19.52%	0.1719		10.55%	0.1390
			Total	26.47%	0.2080		20.14%	0.2161

From Table 12 directional support for Hypotheses 1.1, 2.1, and 3.1 can be found. The mean accuracy for the visual search interface is 28.96% (*standard deviation 0.2240*) while the mean accuracy for the keyword search interface is 17.67% (*standard deviation 0.1840*). This supports Hypothesis 1.1, which states that the visual search interface will produce higher accuracy rates compared to the keyword search interface. Hypotheses 2.1 states that as result set size increases accuracy will decrease. The mean accuracy for the small result set size is 38.45% (*standard deviation 0.2106*), the mean accuracy for the medium result set size is 17.04% (*standard deviation 0.1833*) and the mean accuracy for the large result set size is 15.57% (*standard deviation 0.1635*), supporting Hypothesis 2.1. The mean accuracy for graduate students is 26.47% (*standard deviation 0.2080*), whereas the mean accuracy for undergraduate students is 20.14% (*standard deviation*

0.2161). This supports Hypothesis 3.1, which stated that accuracy rates of more experienced subjects will be higher than accuracy rates of less experienced subjects.

Figures 13a – 13f provide graphical depictions of the hypothesized results and the study results.

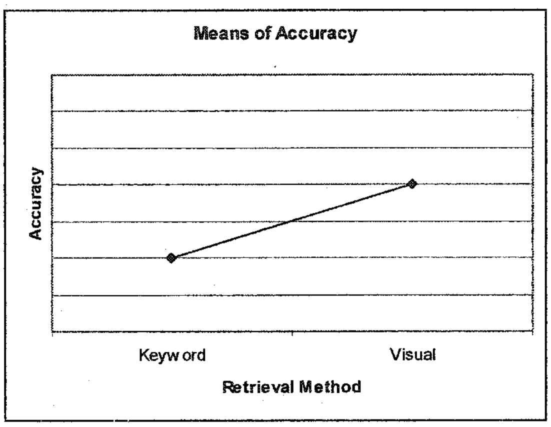


Figure 13a - Search Interface - Accuracy (Hypothesized)

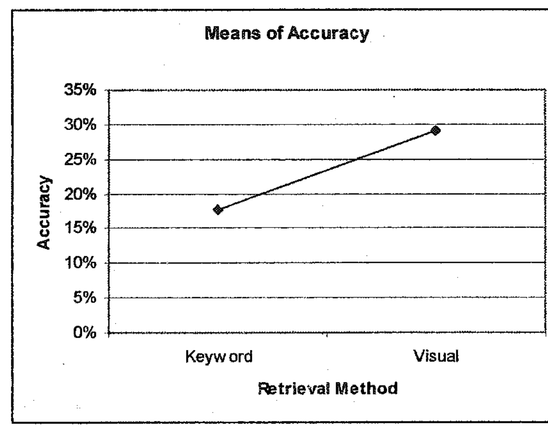


Figure 13b - Search Interface - Accuracy (Actual)

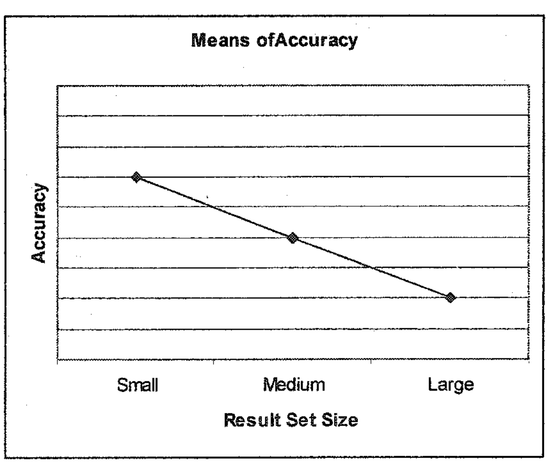


Figure 13c – Result Set Size - Accuracy (Hypothesized)

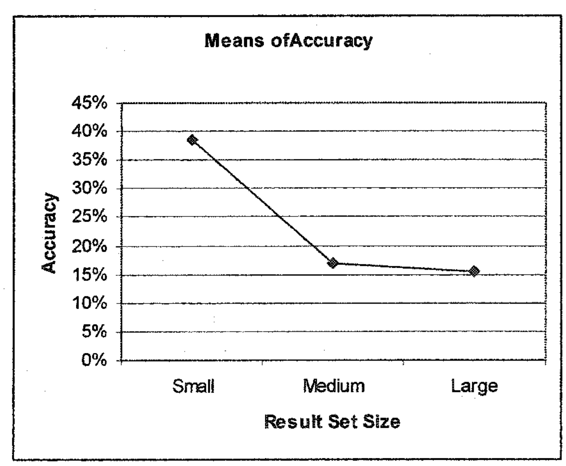


Figure 13d – Result Set Size - Accuracy (Actual)

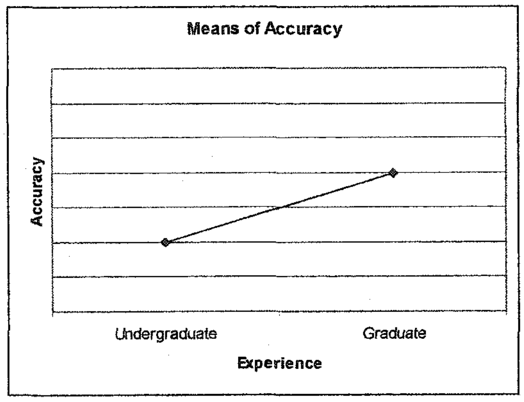


Figure 13e – Education - Accuracy (Hypothesized)

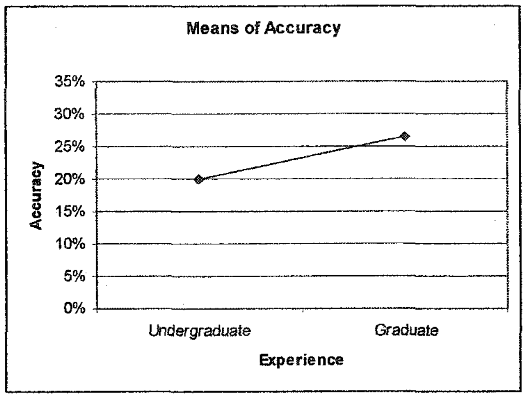


Figure 13f – Education - Accuracy (Actual)

Linear Models for Accuracy.

For further evaluation of accuracy a generalized linear model analysis of variance (GLM ANOVA) was employed using SPSS 12.0 to test for statistical significance of the independent variables (retrieval method, experience) and their interactions for each of the three result set sizes. Because linear models assume homogeneity of variance, statistical tests were run to test for homogeneity of variance. These tests revealed that accuracy, time, and work effort did not meet this condition. To resolve this problem these variables were standardized to the variable with the largest variance. The process for standardization is explained in the next section.

Standardizing Accuracy.

Residuals were calculated for each subject's accuracy score for each of the scenarios. The original residual is equal to the subject's score minus the overall mean for that score. A new residual is then calculated by multiplying the original residual score by the ratio of the standard deviation of the largest accuracy group to the standardization of the original group. A new accuracy score is then computed by adding the new residual score back into the mean score. By following this procedure the means in each group remain as they originally were, whereas their standard deviations become the same.

Table 13 presents the data for the accuracy models after standardizing the data. Because none of the moderating interaction effects were significant a model without interactions is presented in Table 14. Following Table 14, the estimated linear equations are provided.

Table 13

Coefficients for Accuracy with Interactions

Model	BIG Accuracy	MED Accuracy	SML Accuracy		stdBIG Accuracy	stdMED Accuracy	stdSML Accuracy
Intercept	24.160	22.100	47.470		0.368	0.285	0.475
sig.	0.000	0.000	0.000		0.000	0.001	0.000
RetMeth	-10.240	-4.760	-16.610		-0.253	-0.107	-0.166
sig.	0.035	0.385	0.009		0.035	0.385	0.009
Educaction	-9.970	-1.260	-4.900		-0.246	0.028	-0.049
sig.	0.046	0.823	0.447		0.046	0.823	0.447
RetMeth*Educ	2.730	-10.710	4.190		0.067	-0.242	0.042
sig.	0.705	0.195	0.655		0.705	0.195	0.655
R-Squared	0.154	0.121	0.133		0.154	0.121	0.133

Table 14

Coefficients for Accuracy without Interactions

Model	BIG Accuracy	MED Accuracy	SML Accuracy		stdBIG Accurracy	stdMED Accuracy	stdSML Accuracy
Intercept	23.610	24.240	46.630		0.354	0.333	0.466
sig.	0.000	0.000	0.000		0.000	0.000	0.000
RetMeth	-9.030	-9.490	-14.760		-0.223	-0.214	-0.148
sig.	0.013	0.023	0.002		0.013	0.023	0.002
Education	-8.680	-6.300	-2.930		-0.214	-0.142	-0.029
sig.	0.017	0.129	0.529		0.017	0.129	0.529
R-Squared	0.152	0.100	0.130		0.152	0.100	0.130

The full generalized linear equation model with the interaction effect follows the standard linear equation:

$$\hat{Y} = \beta_0 + \beta_1 * \alpha_1 + \beta_2 * \alpha_2 + \beta_3 * \alpha_1\alpha_2 + \varepsilon$$

where:

$\hat{Y}$ is the predicted value of the dependent variable

β_0 is the intercept coefficient

β_1, β_2, β_3 are the independent variable coefficients

α_1, α_2 are the independent variables

$\alpha_1\alpha_2$ is the interaction between the two independent variables

ε is the error term.

$$Accuracy_{Big} = 24.16 + (-10.24)*RetrievalMethod + (-9.97)*Education + 2.73*RetrievalMethod*Education + Error$$

$$Accuracy_{Med} = 22.10 + (-4.76)*RetrievalMethod + (-1.26)*Education + (-4.19)*RetrievalMethod*Education + Error$$

$$Accuracy_{Sml} = 47.47 + (-16.61)*RetrievalMethod + (-4.90)*Education + 4.19*RetrievalMethod*Education + Error$$

None of the interaction effects of search interface and result set size turned out to be significant in the individual or pooled models. Hypotheses 4.1 through 4.4 are not supported. The interactions were removed from the linear equations and a new model was run with only main effects yielding the following equations:

$$Accuracy_{Big} = 23.61 + (-9.03)*RetrievalMethod + (-8.68)*Education + Error$$

$$Accuracy_{Med} = 24.24 + (-9.49)*RetrievalMethod + (-6.30)*Education + Error$$

$$Accuracy_{Big} = 46.63 + (-14.76)*RetrievalMethod + (-2.93)*Education + Error$$

Analysis of Time

Table 15 provides descriptive statistics on the dependent variable time; including time based on result set size, retrieval method, and education.

Table 15

Descriptive Statistics for Time

	Overall	(n=75)		Visual	(n=40)		Keyword	(n=35)
Time	Mean	Std Dev		Mean	Std Dev		Mean	Std Dev
Sml	302.81	229.35		312.15	229.89		292.14	231.61
Med	349.37	225.99		375.60	251.24		319.40	192.38
Big	344.63	274.93		392.20	339.75		290.26	162.50
Total	996.81	244.25		1079.95	277.56		901.80	196.08
				Graduate	(n=42)		Undergrad	(n=33)
			Time	Mean	Std Dev		Mean	Std Dev
			Sml	351.67	237.29		240.64	205.85
			Med	410.38	233.34		271.73	192.96
			Big	425.64	325.54		241.52	139.42
			Total	1187.69	268.54		753.89	180.42

From Table 15 directional support for Hypotheses 1.2 and 2.2 can be found. The mean time (in seconds) for the visual search interface is 360 (*standard deviation 278*) while the mean time for the keyword search interface is 301 (*standard deviation 196*). Hypothesis 1.2, which states that the visual search interface will take longer time compared to the keyword search interface, is supported. Hypotheses 2.2 stated that as result set size increases time will also increase. This hypothesis is supported for the visual search interface only. The mean time for the small result set size is 312 (*standard deviation 230*), the mean time for the medium result set size is 376 (*standard deviation 251*) and the mean time for the large result set size is 392 (*standard deviation 278*). The mean time for the graduate students is 396 (*standard deviation 269*), whereas the mean time for undergraduate students is 251 (*standard deviation 180*). These results are opposite what Hypothesis 3.2 predicted. Hypothesis 3.2 states that the time of more experienced subjects will be lower than the time of less experienced subjects.

Figures 14a – 14f provide graphical depictions of the hypothesized results and the study results.

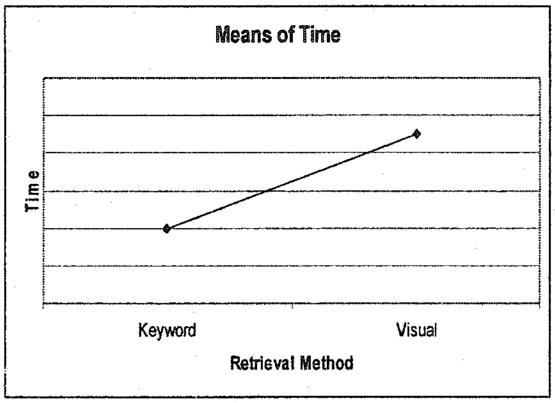


Figure 14a – Search Interface - Time (Hypothesized)

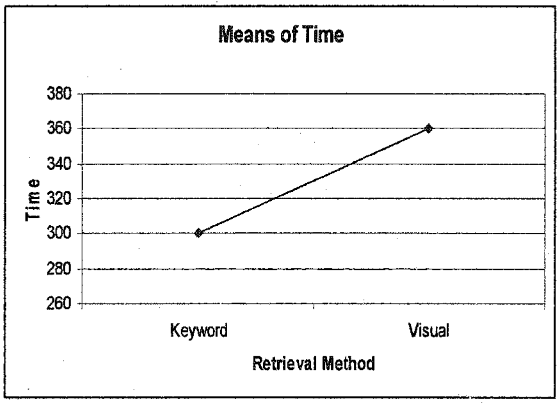


Figure 14b – Search Interface - Time (Actual)

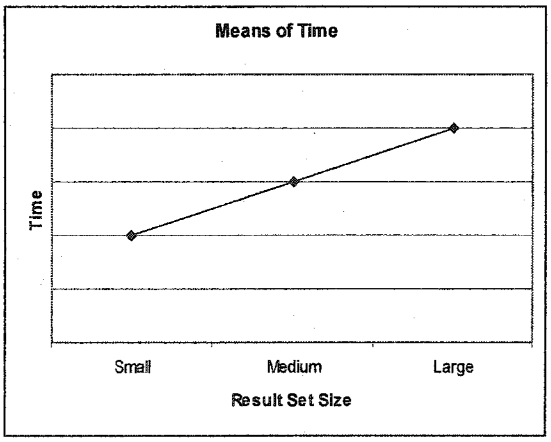


Figure 14c – Result Set Size - Time (Hypothesized)

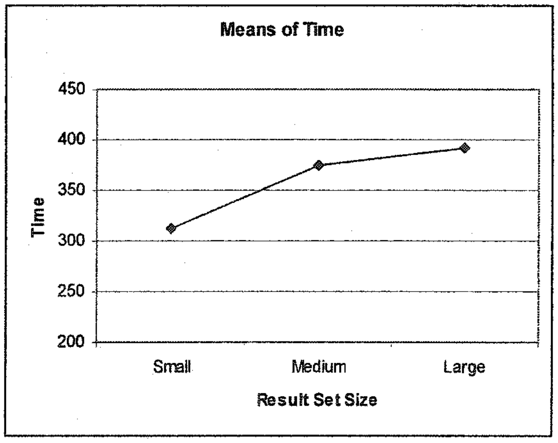


Figure 14d – Result Set Size - Time (Actual)

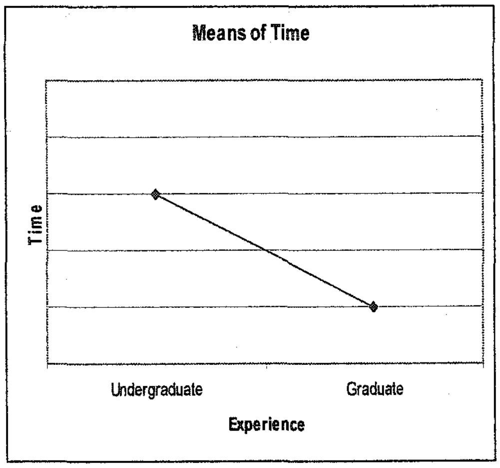


Figure 14e – Education - Time (Hypothesized)

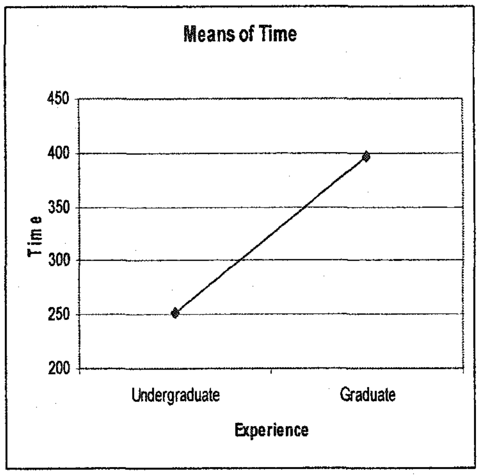


Figure 14f – Education - Time (Actual)

Linear Models for Time.

As with accuracy, time was analyzed via a generalized linear model analysis of variance (GLM ANOVA) to test for statistical significance of the independent variables (retrieval method and education) and their interaction for each of the three result set sizes. Time failed the homogeneity of variance tests so standardized variables, based on the largest variance were created following the same procedure as documented in the *Analysis of Accuracy* section.

Table 16 presents the data for the time models after standardizing the data. Because it was found that none of the moderating interaction effects were significant a model without interactions is presented in Table 17. Following Table 17, the estimated linear equations are provided.

Table 16

Coefficients for Time with Interactions

Model	BIG Time	MED Time	SML Time		stdBIG Time	stdMED Time	stdSML Time
Intercept	491.44	427.70	377.57		491.44	444.660	392.420
sig.	0.000	0.000	0.000		0.000	0.000	0.000
RetMeth	-145.44	-38.28	-57.25		-145.440	-46.560	-68.630
sig.	0.072	0.573	0.416		0.072	0.573	0.416
Education	-233.49	-122.58	-153.92		-233.49	-149.13	-184.51
sig.	0.006	0.083	0.037		0.006	0.083	0.037
RetMeth*Educ	111.56	-30.59	92.29		111.56	-37.22	110.630
sig.	0.357	0.764	0.384		0.357	0.764	0.384
R-Squared	0.153	0.131	0.070		0.153	0.108	0.070

Table 17

Coefficients for Time without Interactions

Model	BIG Time	MED Time	SML Time		stdBIG Time	stdMED Time	stdSML Time
Intercept	469.13	433.81	359.110		469.13	452.10	370.300
sig.	0.000	0.000	0.000		0.000	0.000	0.000
RetMeth	-96.13	-51.80	-16.46		-96.13	-63.02	-19.73
sig.	0.112	0.305	0.754		0.112	0.305	0.754
Education	-181.00	-136.97	-110.50		-181.00	-166.64	-132.46
sig.	0.004	0.008	0.039		0.004	0.008	0.039
R-Squared	0.143	0.107	0.060		0.143	0.107	0.060

The full linear models with the interaction effects follow:

$Time_{Big}$ = 491 + (-145)*RetrievalMethod + (-233)*Education + 112*RetrievalMethod*Education + Error

$$Time_{Med} = 428 + (-38)*RetrievalMethod + (-123)*Education + (-31)*RetrievalMethod*Education + Error$$

$$Time_{Sml} = 378 + (-57)*RetrievalMethod + (-154)*Education + 92*RetrievalMethod*Education + Error$$

Because none of the interactions for time turned out to be significant, in the individual models for size or in the repeated measures model they were removed from the linear equations and a new model was run with only main effects yielding the following equations:

$$Time_{Big} = 469 + (-96)*RetrievalMethod + (-181)*Education + Error$$

$$Time_{Med} = 433 + (-52)*RetrievalMethod + (-137)*Education + Error$$

$$Time_{Big} = 370 + (-20)*RetrievalMethod + (-133)*Education + Error$$

A Note on the Difference between Keyword and Visual Time.

For the purposes of this experiment, the subjects were not aware of the differences in result set size. One would expect the keyword searchers to spend approximately the same amount of time on each scenario. That is, they would typically type in as many keywords as they could think of regardless of the size of the results. They would query until they felt satisfied with their results. This should be a relatively constant time. This assumption held true in this experiment as demonstrated by the fact that time differences between the small and large result set size searches for those using the keyword search interface was two seconds, and the medium search varied only by 19 seconds from the other two. On the other hand, those that used the visual interface needed to navigate increasingly larger tree structures as the result sets grew. Thus, is predicted that more time would be spent searching with the visual interface.

Analysis of Work Effort

Table 18 provides descriptive statistics on the dependent variable work effort including effort based on result set size, retrieval method, and education.

Table 18

Descriptive Statistics for Work Effort

	Overall	(n=70)		Visual	(n=38)		Keyword	(n=32)
Effort	Mean	Std Dev		Mean	Std Dev		Mean	Std Dev
Sml	60.80	15.59		65.40	14.40		55.33	15.38
Med	64.61	15.18		69.14	10.97		59.24	17.73
Big	52.05	15.67		56.78	14.60		46.43	15.23
Total	59.15	16.28		63.77	14.29		53.67	16.87
				Graduate	(n=41)		Undergrad	(n=29)
			Effort	Mean	Std Dev		Mean	Std Dev
			Sml	61.95	12.77		59.17	19.01
			Med	65.24	11.87		63.72	19.10
			Big	54.56	15.36		48.49	15.68
			Total	60.58	14.04		57.13	18.91

Analysis of the data in Table 18 show no support for Hypotheses 1.3, 2.3 and 3.3. The mean work effort for the visual search interface is 63.77 (*standard deviation 14.29*) while the mean work effort for the keyword search interface is 53.67 (*standard deviation 16.87*) opposite of what Hypothesis 1.3 predicts. Hypothesis 1.3 states that the visual search interface will take less effort compared to the keyword search interface. Similarly, Hypotheses 2.3 states that as result set size increases work effort will also increase. This hypothesis is not supported. The mean work effort for the small result set size is 60.80 (*standard deviation 15.59*), the mean work effort for the medium result set size is 64.61 (*standard deviation 15.18*) and the mean work effort for the large result set size is 59.15 (*standard deviation 16.28*). The mean work effort for the graduate students (proxy for high experience) is 60.58 (*standard deviation 14.04*), whereas the mean work effort for undergraduate students (proxy for low experience) is 57.13 (*standard deviation 18.91*). These results are also opposite what Hypothesis 3.3 predicts. Hypothesis 3.3 states that work effort of more experienced subjects will be lower than work effort of less experienced subjects.

Figures 15a – 15f provide graphical depictions of the hypothesized results and the study results.

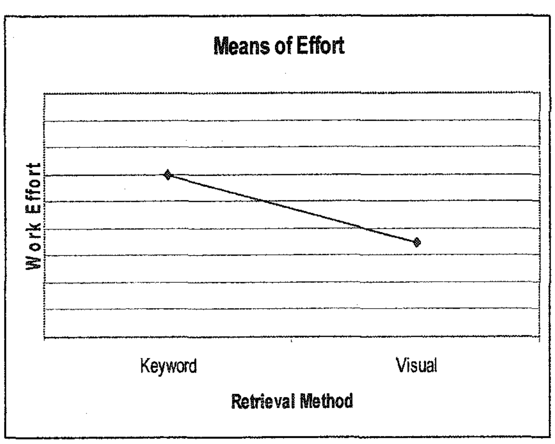


Figure 15a – Search Interface – Effort (Hypothesized)

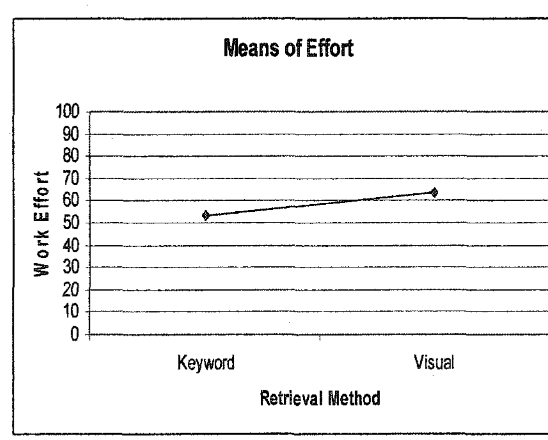


Figure 15b – Search Interface - Effort (Actual)

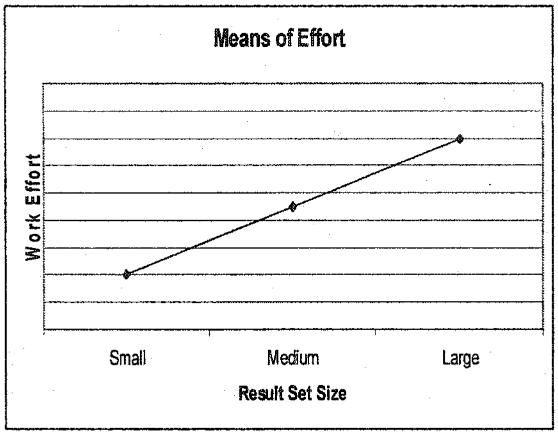


Figure 15c – Result Set Size - Effort (Hypothesized)

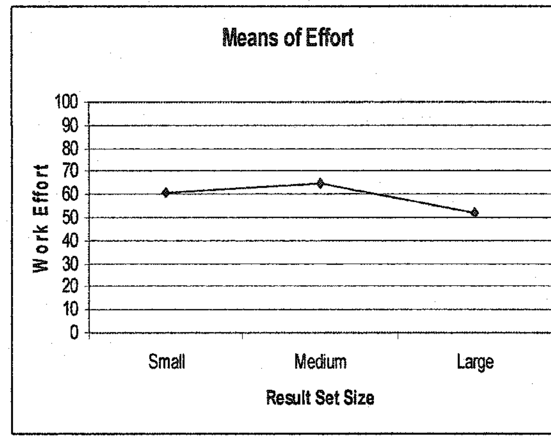


Figure 15d – Result Set Size - Effort (Actual)

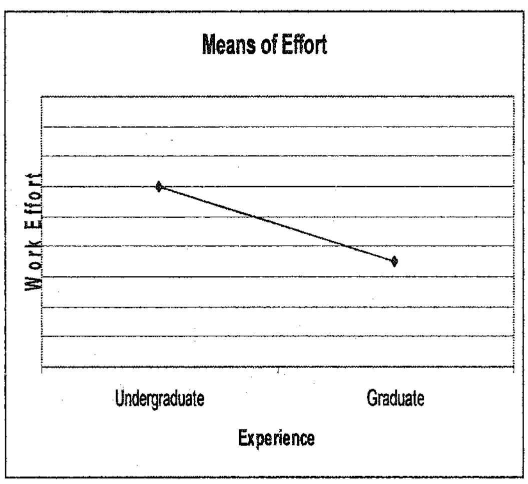


Figure 15e – Education - Effort (Hypothesized)

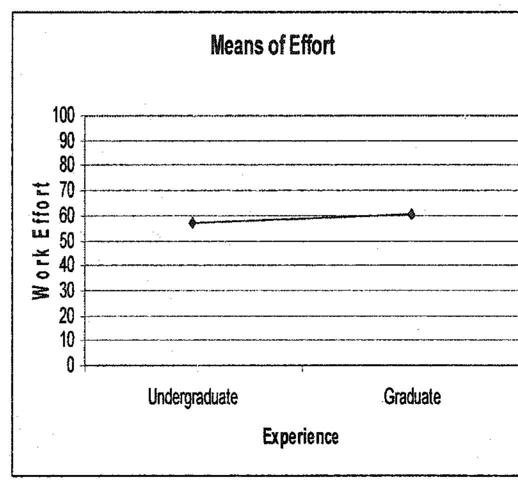


Figure 15f – Education - Effort (Actual)

Linear Models for Work Effort.

As with accuracy and time, work effort was analyzed via a generalized linear model analysis of variance (GLM ANOVA) to test for statistical significance of the independent variables (retrieval method and education) and their interaction for each of the three result set sizes. Because linear models assume homogeneity of variance, statistical tests were run to test for homogeneity of variance. These tests reveal that work effort did not meet this condition. To resolve this problem the work effort variable is standardized to the factor with the largest variance following the same procedure as was documented in the *Analysis of Accuracy* section.

Table 19 presents the data for the work effort models after standardizing the data. Because it was found that none of the moderating interaction effects were significant a model without interactions is presented in Table 20. Following Table 20, the estimated linear equations are provided.

Table 19

Coefficients for Work Effort with Interactions

Model	BIG Effort	MED Effort	SML Effort		stdBIG Effort	stdMED Effort	stdSML Effort
Intercept	58.44	68.83	65.44		58.44	68.97	65.460
sig.	0.000	0.000	0.000		0.000	0.000	0.000
RetMeth	-8.37	-7.75	-7.53		-8.37	-8.00	-7.580
sig.	0.074	0.095	0.112		0.074	0.095	0.112
Education	-3.94	0.73	-0.09		-3.94	0.75	-0.09
sig.	0.419	0.879	0.986		0.419	0.879	0.986
RetMeth*Educ	-5.03	-5.28	-6.26		-5.03	-5.45	-6.300
sig.	0.486	0.460	0.392		0.486	0.460	0.392
R-Squared	0.155	0.117	0.124		0.155	0.117	0.124

Table 20

Coefficients for Work Effort without Interactions

Model	BIG Effort	MED Effort	SML Effort		stdBIG Effort	stdMED Effort	stdSML Effort
Intercept	59.40	69.84	66.640		59.40	70.01	66.670
sig.	0.000	0.000	0.000		0.000	0.000	0.000
RetMeth	-10.45	-9.93	-10.11		-10.45	-10.25	-10.17
sig.	0.004	0.006	0.006		0.004	0.006	0.006
Education	-6.23	-1.67	-2.93		-6.23	-1.72	-2.95
sig.	0.085	0.637	0.420		0.085	0.637	0.420
R-Squared	0.149	0.110	0.114		0.149	0.110	0.114

The full linear model with the interaction effect follows:

$Effort_{Big}$ = 58.44 + (-8.37)*RetrievalMethod + (-3.94)*Education + (-5.03)*RetrievalMethod*Education + Error

$Effort_{Med}$ = 68.83 + (-7.75)*RetrievalMethod + 0.73*Education + (-5.28)*RetrievalMethod*Education + Error

$Effort_{Sml}$ = 65.44 + (-7.53)*RetrievalMethod + (-0.09)*Education + (-6.26)*RetrievalMethod*Education + Error

Because none of the interactions turned out to be significant, in the individual or pooled models for work effort they were removed from the model and new linear equations were calculated using only the main effects yielding the following equations:

$Effort_{Big}$ = 59.40 + (-10.45)*RetrievalMethod + (-6.23)*Education + Error

$Effort_{Med}$ = 69.84+ (-9.93)*RetrievalMethod + (-1.67)*Education + Error

$Effort_{Sml}$ = 66.64 + (-10.11)*RetrievalMethod + (-2.95)*Education + Error

Analysis of Satisfaction

Table 21 provides descriptive statistics on the dependent variable satisfaction including satisfaction based on result set size, retrieval method, and education.

Table 21

Descriptive Statistics for Satisfaction

	Overall	(n=73)		Visual	(n=38)		Keyword	(n=35)
Satisfaction	Mean	Std Dev		Mean	Std Dev		Mean	Std Dev
Sml	15.08	3.43		14.85	3.06		15.33	3.82
Med	15.02	3.37		14.72	3.11		15.34	3.65
Big	15.53	3.63		14.96	3.76		16.15	3.43
Total	15.21	3.47		14.84	3.30		15.61	3.62
				Graduate	(n=41)		Undergrad	(n=32)
			Satisfaction	Mean	Std Dev		Mean	Std Dev
			Sml	14.87	3.24		15.35	3.69
			Med	15.37	3.36		14.55	3.37
			Big	15.63	3.03		15.40	4.32
			Total	15.29	3.21		15.10	3.80

From Table 21 no support for Hypotheses 1.4 and 2.4 can be found. The mean of satisfaction for the visual search interface is 14.84 (*standard deviation 3.30*) while the mean of satisfaction for the keyword search interface is 15.61 (*standard deviation 3.62*). Hypothesis 1.4, which states that the visual search interface will be more satisfying compared to the keyword search interface, is not supported. Hypotheses 2.4 states that as result set size increases satisfaction will decrease. This hypothesis also is not supported.

The mean of satisfaction for the small result set size is 15.08 (*standard deviation 3.43*), the mean of satisfaction for the medium result set size is 15.02 (*standard deviation 3.37*) and the mean of satisfaction for the large result set size is 15.21 (*standard deviation 3.63*). The mean of satisfaction for the graduate students is 15.29 (*standard deviation 3.21*), whereas the mean of satisfaction for undergraduate students is 15.10 (*standard deviation 3.8*). These results show directional support for Hypothesis 3.4. Hypothesis 3.4 states that satisfaction of more experienced subjects will be higher than satisfaction of less experienced subjects.

Figures 16a – 16f provide graphical depictions of the hypothesized results and the study results.

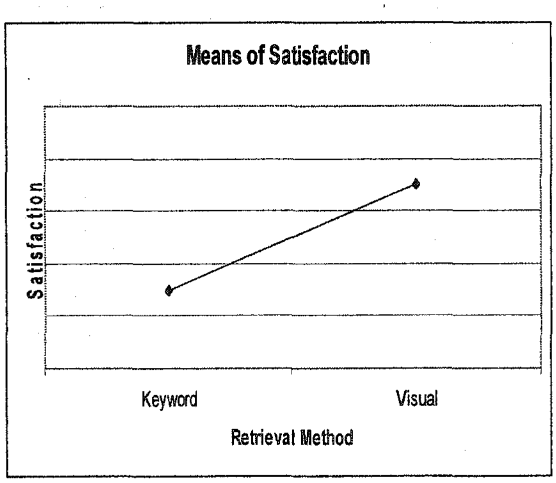


Figure 16a – Search Interface – Satisfaction (Hypothesized)

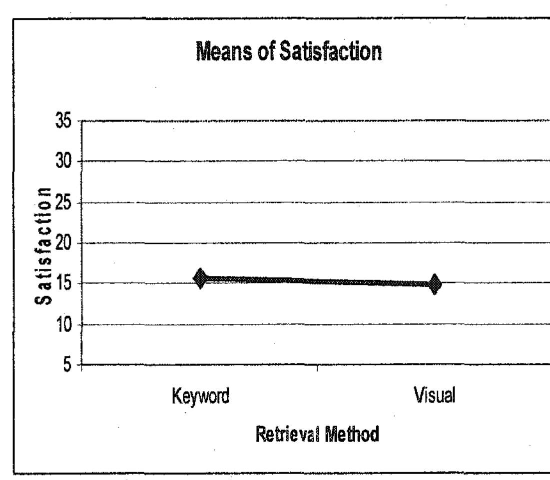


Figure 16b – Search Interface - Satisfaction (Actual)

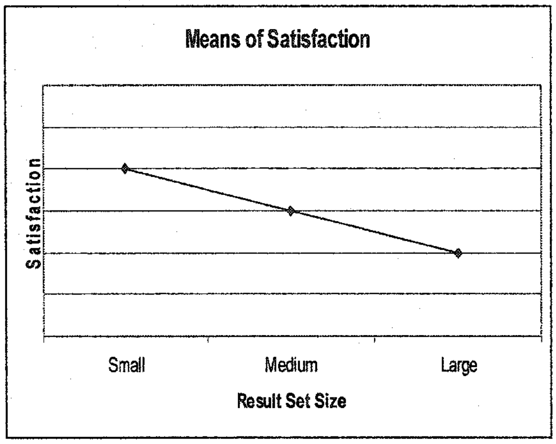


Figure 16c – Result Set Size - Satisfaction (Hypothesized)

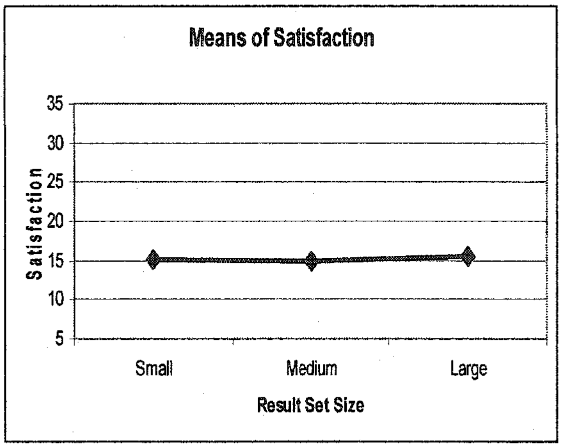


Figure 16d – Result Set Size - Satisfaction (Actual)

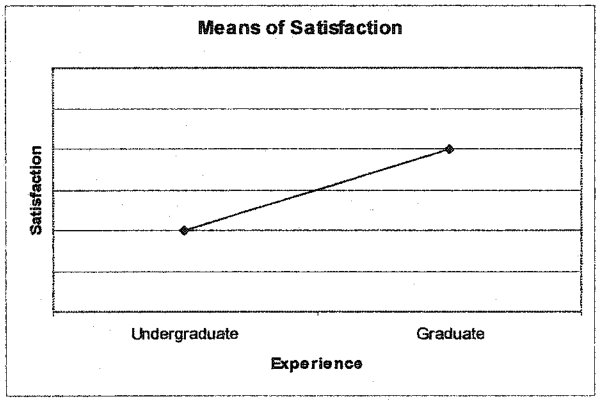


Figure 15e – Education - Satisfaction (Hypothesized)

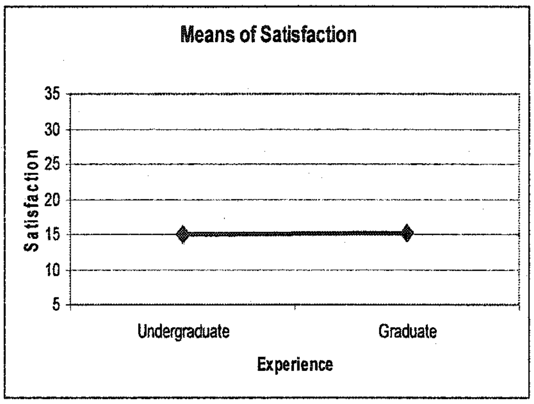


Figure 15f – Education - Satisfaction (Actual)

Linear Models for Satisfaction.

As with accuracy, time, and work effort, satisfaction was analyzed via a generalized linear model analysis of variance (GLM ANOVA) to test for statistical significance of the independent variables (retrieval method and education) and their interaction for each of the three result set sizes. Because linear models assume homogeneity of variance, statistical tests were run to test for homogeneity of variance.

These tests revealed that satisfaction did meet this condition; however for consistency's sake standardized variables were still formulated. Examining the estimated linear coefficients showed little to no change between the standardized and non-standardized variables.

Table 22 presents the data for the work effort models after standardizing the data. Because it was found that none of the moderating interaction effects were significant a model without interactions is presented in Table 23. Following Table 23, the estimated linear equations are provided.

Table 22

Coefficients for Satisfaction with Interactions

Model	BIG Satisfact	MED Satisfact	SML Satisfact		stdBIG Satisfact	stdMED Satisfact	stdSML Satisfact
Intercept	14.86	14.80	14.75		14.86	14.78	14.73
sig.	0.000	0.000	0.000		0.000	0.000	0.000
RetMeth	1.66	1.26	0.25		1.66	1.36	0.26
sig.	0.150	0.239	0.820		0.150	0.239	0.820
Education	0.23	-0.17	0.23		0.23	-0.18	0.25
sig.	0.848	0.878	0.839		0.848	0.878	0.839
RetMeth*Educ	-1.05	-1.40	0.48		-1.05	1.50	0.510
Sig.	0.545	0.384	0.770		0.545	0.384	0.770
R-Squared	0.034	0.035	0.011		0.034	0.035	0.011

Table 23

Coefficients for Satisfaction without Interactions

Model	BIG Satisfact.	MED Satisfact.	SML Satisfact.		stdBIG Satisfact.	stdMED Satisfact.	stdSML Satisfact.
Intercept	15.08	15.08	14.650		15.08	15.09	14.630
sig.	0.000	0.000	0.000		0.000	0.000	0.000
RetMeth	1.20	0.64	0.46		1.20	0.69	0.49
sig.	0.162	0.419	0.570		0.162	0.419	0.570
Education	-0.28	-0.85	0.47		-0.28	-0.91	0.50
sig.	0.747	0.291	0.568		0.747	0.291	0.568
R-Squared	0.029	0.024	0.010		0.029	0.024	0.010

The full linear equations with the interaction effects follow:

$Satisfaction_{Big}$ = 14.86 + 1.66*RetrievalMethod + 0.23*Education + (-1.05)*RetrievalMethod*Education + Error

$Satisfaction_{Med}$ = 14.80 + 1.26*RetrievalMethod + (-0.17)*Education + (-1.40)*RetrievalMethod*Education + Error

$Satisfaction_{Sml}$ = 14.75 + 0.25*RetrievalMethod + 0.23*Education + 0.48*RetrievalMethod*Education + Error

Because none of the interactions turned out to be significant, in the individual or pooled models for search interface and result set size, the interactions were removed from the model and new linear equations were run with only the main effects yielding the following equations:

$Satisfaction_{Big}$ = 15.08 + 1.20*RetrievalMethod + (-0.28)*Education + Error

$Satisfaction_{Med}$ = 15.09 + 0.69*RetrievalMethod + (-0.85)*Education + Error

$$Satisfaction_{Sml} = 14.65 + 0.46*RetrievalMethod + 0.47*Education + Error$$

Pooled Response Analysis

To measure the effect of size, two different statistical analyses were performed. The first analysis employed a pooled response analysis for each of the dependent variables, with an additional categorical (dummy) variable created to distinguish size. When pooling the responses two dummy variables are created to distinguish between the different result set sizes (small, medium, and large). Table 24 presents the estimated linear coefficients and their significance (p-values) for each of the four factors (accuracy, time, work effort, and satisfaction) that make up the dependent variable retrieval effectiveness. These statistics were calculated using the univariate generalized linear modeling statistical procedures within SPSS 12.0. Because the interactions are not statistically significant another model was run on the pooled data set without the interactions. Those results are shown in Table 25. Finally, estimated linear equations are provided for the models following each table.

Table 24

Pooled Sample Coefficients with Interactions

Model	Pooled Accuracy w/Interactions	Pooled Time w/Interactions	Pooled Effort w/Interactions	Pooled Satisfaction w/ Interactions
Intercept	47.49	384	65.87	14.81
sig.	0.000	0.000	0.000	0.000
RetMeth	-15.73	-41	-7.84	0.78
sig.	0.000	0.501	0.046	0.400
Education	-4.96	-170	-1.1	0.10
sig.	0.066	0.000	0.693	0.882
SizeBig	-25.46	80	-8.62	0.11
sig.	0.000	0.127	0.011	0.889
SizeMed	-45.17	63	3.74	-0.13
sig.	0.000	0.226	0.267	0.876
RetMeth*Educ	2.27	58	-5.52	-0.65
sig.	0.563	0.359	0.181	0.493
RetMeth*Big	5.55	-82	-0.28	0.71
sig.	0.246	0.285	0.955	0.539
RetMeth*Med	14.76	-36	0.17	0.13
sig.	0.002	0.636	0.973	0.901
R-Squared	0.575	0.115	0.220	0.022

Table 24 shows that none of the interaction effects are significant, thus there is no support for Hypotheses 4.1 – 4.4 which address experience moderating search interface. Furthermore, Hypotheses 5.1 – 5.4, which address the moderating role of result set size on search interface to the factors of the dependent variable retrieval effectiveness, are likewise not supported.

From the pooled data with interactions the following linear equations can be derived:

$Accuracy_{pooled}$ = 47.49 + (-15.73)*RetMeth + (-4.96)*Education + (-25.46)*SizeBig + (-45.17)*SizeMed + 2.27*RetMeth*Education + 5.55*RetMeth*SizeBig + 14.76*RetMeth*SizeMed + Error

$Time_{pooled}$ = 384 + (-41)*RetMeth + (-170)*Education + 80*SizeBig + 63*SizeMed + 58*RetMeth*Education + (-82)*RetMeth*SizeBig + (-36)*RetMeth*SizeMed + Error

$WorkEffort_{pooled}$ = 65.87 + (-7.84)*RetMeth + (-1.10)*Education + (-8.62)*SizeBig + 3.74*SizeMed + (-5.52)*RetMeth*Education + (-0.28)*RetMeth*SizeBig + 0.17*RetMeth*SizeMed + Error

$Satisfaction_{pooled}$ = 14.81 + 0.78*RetMeth + 0.10*Education + 0.11*SizeBig + (-0.13)*SizeMed + (-0.65)*RetMeth*Education + 0.71*RetMeth*SizeBig + 0.13*RetMeth*SizeMed + Error

Table 25

Pooled Sample Coefficients without Interactions

Model	Pooled Accuracy w/o Interactions	Pooled Time w/o Interactions	Pooled Effort w/o Interactions	Pooled Satisfaction w/o Interactions
Intercept	43.88	391	66.94	14.81
sig.	0.000	0.000	0.000	0.000
RetMeth	-7.96	0.55	-10.16	0.77
sig.	0.000	0.080	0.000	0.103
Education	-3.89	-143	-3.61	-0.22
sig.	0.051	0.000	0.079	0.645
SizeBig	-22.88	42	-8.75	0.45
sig.	0.000	0.273	0.000	0.431
SizeMed	-38.28	47	3.80	-0.06
sig.	0.000	0.222	0.123	0.915
R-Squared	0.556	0.107	0.213	0.017

Removing the interactions from the model produces the results in Table 25. From the pooled data without interactions the following linear equations can be derived:

$$Accuracy_{pooled} = 43.88 + (-7.96)*RetMeth + (-3.89)*Education + (-22.88)*SizeBig + (-38.28)*SizeMed + Error$$

$$Time_{pooled} = 391 + 0.55*RetMeth + (-143)*Education + 42*SizeBig + 47*SizeMed + Error$$

$$WorkEffort_{pooled} = 66.94 + (-10.16)*RetMeth + (-3.61)*Education + (-8.75)*SizeBig + 3.80*SizeMed + Error$$

$$Satisfaction_{pooled} = 14.81 + 0.77*RetMeth + (-0.22)*Education + 0.45*SizeBig + (-0.06)*SizeMed + Error$$

By incorporating size into the pooled model we see the R^2 multiple increases for depended variable accuracy from the 12-15% range to 55-58% range depending on the model. Examining the time dependent variable by incorporating size into the model, the R^2 multiple stabilizes at 11% from a range of 7-15% in the individual models. Similar to accuracy, the R^2 multiple for work effort increases from 11-16% in the individual models, to 21-22% for the pooled model. In the case of satisfaction, very little changes. Like time, the pooled model R^2 multiple for satisfaction stabilizes in the middle of the range of the individual R^2 multiple for satisfaction. The individual R^2 multiple for satisfaction ranges from 1-3.5%, whereas the pooled models have an R^2 multiple of approximately 2%.

Repeated Measures Analysis

By pooling the data together there is a chance for an incorrect reporting of the estimated coefficients and R^2 multiple, based on an incorrect error variance. When employing a repeated measures design there are actual two error terms that need to be taken into account. The first term is the overall error term, and the second term takes into account the error for the within subject portion of the experimental design. To avoid this problem and to test for significance a GLM ANOVA repeated measures statistical analysis (Girdon, 1992) was performed using NCSS 2004 (Hintz, 2004). NCSS 2004 was chosen over SPSS 12.0 for its flexibility in specifying models for repeated measures. *F* statistics and significance (*p*-values) are given for two models, the original model with interactions and the adjusted model without the nonsignificant interaction, for each of the dependent variables. Both original and standardized data analysis is presented:

Table 26

Repeated Measures Analysis for Accuracy

		Accuracy w/Interactions		stdAccuracy w/Interactions	
Term	DF	F-Ratio	Prob. Level	F-Ratio	Prob. Level
A: RetMeth	1	10.65	0.002	9.58	0.003
B: Education	1	3.08	0.084	4.15	0.045
AB	1	0.03	0.854	0.12	0.730
C(AB): SubID	71	4.28	0.000	3.72	0.000
D: Size	2	79.68	0.000	20.74	0.000
AD	2	1.18	0.309	0.6	0.550
		Accuracy w/o Interactions		stdAccuracy w/o Interactions	
	DF	F-Ratio	Prob. Level	F-Ratio	Prob. Level
A: RetMeth	1	10.79	0.002	9.58	0.003
B: Education	1	3.09	0.083	4.13	0.046
C(AB): SubID	72	4.21	0.000	3.69	0.000
D: Size	2	81.14	0.000	20.48	0.000

From Table 26, examining the standardized scores without interactions, we find significance for all three of our independent variables (retrieval method, education (proxied by education), and result set size) to the dependent factor accuracy. Furthermore, no significance for any of the interaction effects is found. This analysis matches the findings from the individual and pooled GLM equation models presented earlier in this chapter.

Table 27

Repeated Measures Analysis for Time

		Time w/Interactions		stdTime w/Interactions	
Term	DF	F-Ratio	Prob. Level	F-Ratio	Prob. Level
A: RetMeth	1	2.45	0.122	2.28	0.136
B: Education	1	18.45	0.000	18.17	0.000
AB	1	0.77	0.382	0.69	0.409
C(AB): SubID	71	1.14	0.246	1.15	0.233
D: Size	2	0.87	0.423	0.68	0.508
AD	2	0.6	0.548	0.43	0.649
		Time w/o Interactions		stdTime w/o Interactions	
Term	DF	F-Ratio	Prob. Level	F-Ratio	Prob. Level
A: RetMeth	1	2.83	0.097	2.63	0.109
B: Education	1	19.02	0.000	18.73	0.000
C(AB): SubID	72	1.15	0.241	1.16	0.226
D: Size	2	0.95	0.389	0.75	0.473

From Table 27 we find significance only for the independent variable education to the dependent factor time. This significance is opposite what was hypothesized. Furthermore, no significance for any of the interaction effects is found. Further detailed analysis does find significance for search interface when only the visual search interface subjects are analyzed. Again, this is assumed because the keyword searchers are predicted to spend the same amount of time on any search task whereas those using visual interfaces will take more time the larger the result set size.

Table 28

Repeated Measures Analysis for Work Effort

		Effort w/Interactions		stdEffort w/Interactions	
Term	DF	F-Ratio	Prob. Level	F-Ratio	Prob. Level
A: RetMeth	1	11.85	0.001	11.84	0.00
B: Education	1	1.56	0.216	1.54	0.22
AB	1	0.8	0.375	0.8	0.38
C(AB): SubID	66	5.78	0.000	5.79	0.00
D: Size	2	34.64	0.000	33.78	0.00
AD	2	0.01	0.989	0	1.00
		Effort w/o Interactions		stdEffort w/o Interactions	
Term	DF	F-Ratio	Prob. Level	F-Ratio	Prob. Level
A: RetMeth	1	11.17	0.001	11.17	0.001
B: Education	1	1.38	0.245	1.36	0.247
C(AB): SubID	67	5.85	0.000	5.86	0.000
D: Size	2	35.29	0.000	34.49	0.000

From Table 28 we find significance for two of the three independent variables (retrieval method and result set size) to the dependent factor work effort. No significance is found for education nor is significance found for either of the interaction effects. Again, this analysis matches the findings from the GLM equation models presented earlier in this chapter.

Table 29

Repeated Measures Analysis for Satisfaction

		Satisfaction w/Interactions		stdSatisfaction w/Interactions	
Term	DF	F-Ratio	Prob. Level	F-Ratio	Prob. Level
A: RetMeth	1	1.05	0.308	1.03	0.314
B: Education	1	0.1	0.749	0.11	0.745
AB	1	0.21	0.647	0.21	0.648
C(AB): SubID	69	5.59	0.000	5.62	0.000
D: Size	2	1.25	0.290	1.15	0.321
AD	2	0.54	0.586	0.45	0.641
		Satisfaction w/o Interactions		stdSatisfaction w/o Interactions	
Term	DF	F-Ratio	Prob. Level	F-Ratio	Prob. Level
A: RetMeth	1	1.2	0.277	1.17	0.282
B: Education	1	0.1	0.758	0.1	0.755
C(AB): SubID	70	5.57	0.000	5.6	0.000
D: Size	2	1.19	0.306	1.1	0.335

Table 29 shows no significance for any main effect or interaction effect on satisfaction. This analysis also matches the findings from the GLM equation models presented earlier in this chapter.

Correlations among Dependent Variable Factors

Correlation analysis among the dependent variable factors was also performed for additional insights. Table 30 presents the correlation matrix for each of the three measures (small, medium, and large) for each of the four dependent variable factors (accuracy, time, work effort, and satisfaction).

Table 30

Correlation Matrix for Dependent Variable Factors

		Big Acc	Med Acc	Sml Acc	Big Time	Med Time	Sml Time	Big Effort	Med Effort	Sml Effort	Big Sat	Med Sat	Sml Sat
Big Acc	Correlation	1.000	0.612	0.601	0.420	0.076	0.011	0.132	0.129	0.047	-0.189	-0.194	-0.156
	Sig.	.	0.000	0.000	0.000	0.269	0.463	0.141	0.147	0.351	0.061	0.056	0.102
Med Acc	Correlation	0.612	1.000	0.606	0.226	0.203	0.020	0.105	0.086	0.060	-0.099	-0.126	0.108
	Sig.	0.000	.	0.000	0.032	0.049	0.437	0.198	0.243	0.313	0.211	0.154	0.190
Sml Acc	Correlation	0.601	0.606	1.000	0.280	0.053	-0.032	0.244	0.179	0.175	-0.098	-0.015	0.016
	Sig.	0.000	0.000	.	0.010	0.335	0.397	0.022	0.072	0.077	0.214	0.452	0.448
Big Time	Correlation	0.420	0.226	0.280	1.000	0.166	0.102	0.109	0.044	0.120	-0.022	0.034	0.008
	Sig.	0.000	0.032	0.010	.	0.088	0.204	0.189	0.360	0.165	0.429	0.391	0.474
Med Time	Correlation	0.076	0.203	0.053	0.166	1.000	0.125	0.103	0.208	0.088	-0.020	-0.054	0.014
	Sig.	0.269	0.049	0.335	0.088	.	0.154	0.201	0.044	0.238	0.435	0.331	0.455
Sml Time	Correlation	0.011	0.020	-0.032	0.102	0.125	1.000	0.108	0.125	0.333	-0.132	-0.112	0.004
	Sig.	0.463	0.437	0.397	0.204	0.154	.	0.191	0.155	0.003	0.141	0.181	0.487
Big Effort	Correlation	0.132	0.105	0.244	0.109	0.103	0.108	1.000	0.641	0.632	-0.262	-0.122	-0.198
	Sig.	0.141	0.198	0.022	0.189	0.201	0.191	.	0.000	0.000	0.015	0.161	0.053
Med Effort	Correlation	0.129	0.086	0.179	0.044	0.208	0.125	0.641	1.000	0.696	-0.444	-0.294	-0.280
	Sig.	0.147	0.243	0.072	0.360	0.044	0.155	0.000	.	0.000	0.000	0.007	0.010
Sml Effort	Correlation	0.047	0.060	0.175	0.120	0.088	0.333	0.632	0.696	1.000	-0.330	-0.155	-0.139
	Sig.	0.351	0.313	0.077	0.165	0.238	0.003	0.000	0.000	.	0.003	0.104	0.129
Big Sat	Correlation	-0.189	-0.099	-0.098	-0.022	-0.020	-0.132	-0.262	-0.444	-0.330	1.000	0.518	0.587
	Sig.	0.061	0.211	0.214	0.429	0.435	0.141	0.015	0.000	0.003	.	0.000	0.000
Med Sat	Correlation	-0.194	-0.126	-0.015	0.034	-0.054	-0.112	-0.122	-0.294	-0.155	0.518	1.000	0.612
	Sig.	0.056	0.154	0.452	0.391	0.331	0.181	0.161	0.007	0.104	0.000	.	0.000
Sml Sat	Correlation	-0.156	0.108	0.016	0.008	0.014	0.004	-0.198	-0.280	-0.139	0.587	0.612	1.000
	Sig.	0.102	0.190	0.448	0.474	0.455	0.487	0.053	0.010	0.129	0.000	0.000	.

Table 30 shows significant correlations between each of the three accuracy measures (BigAcc, MedAcc, SmlAcc), the three work effort scores (BigEffort, MedEffort, SmlEffort), and the three satisfaction measures (BigSat, MedSat, SmlSat) as would be expected. This suggests that those that were most accurate in the big scenario were also likely to be the most accuracy in the medium and small scenarios. Similarly, those that put forth the most work effort for the big scenario were most likely to put forth the most effort for the medium and small scenarios. Satisfaction results suggest the same pattern. Interestingly, the three time measures (BigTime, MedTime, SmlTime) did not correlate, meaning that those that spent the most time on one scenario were not necessarily the ones that spend the most time on the other scenarios.

Work effort and time are significantly negatively correlated, but their explanatory power of the variance was quite small, ranging from 1-11%. Work effort and satisfaction are also significantly correlated, but again there is very little explanatory power, ranging only from 2-7%.

Summary of Results

Table 31 presents a summary of the results for each of the hypotheses. Following in Chapter 6 a final modified research model is presented as part of the discussion along with implications, limitations, avenues for further research, and conclusionary remarks.

Table 31

Summary of Results

Hypothesis 1: Search Interface to Retrieval Effectiveness – Main Effect	
H1.1: Accuracy	Supported
H1.2: Time	Supported
H1.3: Work Effort	Not Supported – significant in opposite direction
H1.4: Satisfaction	Not Supported
Hypothesis 2: Result set Size to Retrieval Effectiveness – Main Effect	
H2.1: Accuracy	Supported
H2.2: Time	Directional – for the Visual search interface only
H2.3: Work Effort	Not Supported
H2.4: Satisfaction	Not Supported
Hypothesis 3: Experience to Retrieval Effectiveness – Main Effect	
H3.1: Accuracy	Supported
H3.2: Time	Not Supported
H3.3: Work Effort	Not Supported
H3.4: Satisfaction	Directional
Hypothesis 4: Search Interface and Result set Size to Retrieval Effectiveness – Interaction Effect	
H4.1: Accuracy	Not Supported
H4.2: Time	Not Supported
H4.3: Work Effort	Not Supported
H4.4: Satisfaction	Not Supported
Hypothesis 5: Search Interface and Experience to Retrieval Effectiveness – Interaction Effect	
H5.1: Accuracy	Not Supported
H5.2: Time	Not Supported
H5.3: Work Effort	Not Supported
H5.4: Satisfaction	Not Supported

CHAPTER 6

DISCUSSION AND CONCLUSION

This chapter begins with a discussion of the findings from the previous chapter. Implications for both academics and practitioners are then presented. Next, several limitations of the study are discussed. Additional avenues for future research are presented and the work closes with some concluding remarks.

Discussion

Information retrieval from knowledge management systems is an important area of research (Alavi & Liedner, 1999; Markus, 2001). Unfortunately, today most knowledge management systems rely solely on search technology (typically keyword search interfaces) developed primarily for the retrieval of data. Gorla and Walker (1998) and LaBrie and St. Louis (2003) have shown that there are inherent problems with keywords and keyword searching. Knowledge management ontologies (Edgington et al., in press) are one method attempting to address these issues. This study has presented an alternative method to keyword search limitations by demonstrating the superiority of a cognitive based, visual search mechanism. This research asks the question, "*Does the cognitive loading of search mechanisms impact the effectiveness of knowledge retrieval?*"

This experiment attempts to answer this question by comparing a visual tree-view hierarchy search interface with a traditional text-based keyword search interface, in a familiar knowledge management setting of a document management system. It was found that more accurate results are returned with a visual search interface. These results showed over 50% gain in retrieval accuracy over a traditional keyword search mechanism. Furthermore, it was correctly predicted that the visual search interface took subjects more time than the keyword search interface. The time difference however, was not a large amount. In fact, the visual searchers averaged approximately one minute longer than the keyword searchers. Work effort was slightly higher for those using the visual search interface over the keyword search interface. Finally, satisfaction, while hypothesized to be higher for the visual interface turned out to be virtually identical between the two interfaces. With no difference in satisfaction between the two systems, slightly more work effort it would seem plausible that many would be very interested in improving their knowledge workers retrieval accuracy rates by over 50% for approximately one more minute of their time per search.

Due to the replacement of education for experience and because no interaction effects were found to be significant, Figure 17 presents a revised research model that more accurately depicts what was tested and found during this research endeavor.

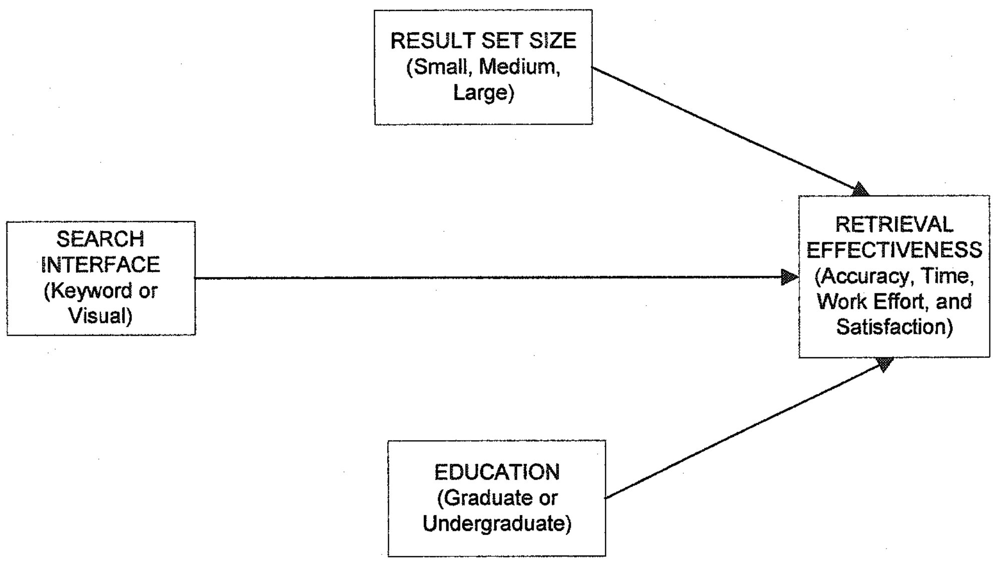


Figure 17. The revised retrieval effectiveness research model.

The lack of significance for satisfaction was unexpected. Results from the pilot study showed a significant difference (2.67 to 5.46 on a 7 point scale) in favor of the visual search interface. Some differences between the administration of the pilot and the experiment may have played a factor in this change. First, doctoral students were used, rather than masters or undergraduate students. Perhaps their knowledge and/or experience with the content of the knowledge base had some bearing on the satisfaction results. Second, the measure for satisfaction changed between the pilot and the experiment. During the pilot only two questions were asked for each scenario – one on satisfaction with the process, one on satisfaction with the results (both 7-point Likert scales). In the experiment the full 12 question EUCS survey instrument (Doll & Torkzadeh, 1988) was utilized. Finally, changes were made on the keyword search interface from the pilot to the experiment. In the pilot the subjects had to go to a separate filter page to run the searches and then return to a main page to select their results. In the experiment the filter page was eliminated and searches were performed on the same page on which the results were displayed. Given all these changes between pilot and experiment administration it is not surprising that some level of satisfaction changed, just how much though was truly unexpected.

However, despite these results the findings have interesting implications. These findings suggest that though those subjects that were presented with the visual interface had to work harder (as measured by the work effort) and longer, they were no *less* satisfied *and* they produced superior accuracy results.

Research Contributions.

This study adds to the body of empirical studies that can be found in the information systems literature that apply cognitive psychology theories in a computerized setting. The results from this study strongly support the theory that visual hierarchical information (recognition) retrieval is more accurate than keyword-based (recall) information retrieval. Results showed a 40-50% increase in accuracy rates using the visual search interface.

This study also validates popular theory surrounding the effort versus accuracy debate popular in IS research. In this study, while accuracy was significantly higher for the visual search interface, it did not come without a price. More time and more effort were spent on the visual search interface over the keyword search interface.

Furthermore, while not statistically significant, satisfaction was slightly higher with the traditional keyword search interface. Results from this study suggest that people continue to be more content sacrificing accuracy for less work effort.

In this study a document management systems was used as a representative type of a knowledge-based management system. KMS documents represent the explicit knowledge captured, structured and codified in a manner consistent with many knowledge management systems. This specific representation of a knowledge management system only deals with explicit knowledge. Other varieties of KMS must deal with implicit, or tacit, knowledge. These types of KMS may need to be operationalized in different manners. While this is beyond the scope of this dissertation, further studies should investigate if a visual search interface would prove as effective for implicit or tacit knowledge, in terms of accuracy, as it does for this instance of explicit knowledge.

Implications for Practice.

The crux of many knowledge management initiatives within organizations is its use. That use depends largely on the effectiveness of its retrieval capabilities. For companies looking for more effective ways of searching, this investigation provides some noteworthy findings. For similar user satisfaction, slightly more work effort, and slightly more time you can greatly improve your retrieval accuracy. The results of this experiment suggest that a 16% increase in time (or about one minute per search scenario) will increase the retrieval accuracy 50-60%. Furthermore, this retrieval accuracy increase can easily be incorporated into today's knowledge management systems with readily available, low cost software (LaBrie & St. Louis, 2004). The benefit of this form of search interface is that it can be added as a supplemental search component, rather than a complete replacement of the traditional keyword search mechanism.

By more accurately retrieving objects from a knowledge repository, for example a knowledge management system for employee skills or, as in the case of this experiment, a document retrieval system, knowledge workers will save time and potentially produce better decisions based on the information found. This, in turn, should result in cost savings for the company. This potential cost savings can then be turned into a return on investment for the IT department and its knowledge management systems (Freeze et al., 2004).

Limitations of the Study

Experimental research inherently has limitations; this section addresses some of the limitations of this study and presents what was done to mitigate those limitations. First, the experiment was conducted in an academic environment that included undergraduate and graduate students. While the MBA students were specifically targeted as a group that would best represent typical knowledge workers, the use of undergraduates was necessary in order to increase the number of subjects. Analysis of the data collected showed a significant difference between these two groups of subjects. So significant, in fact, that the model was changed to reflect education rather than experience. Administering the experiment to additional graduate students might show more significance for some of the other factors other than accuracy. Furthermore, executing this experiment in a business environment could add further external validity. The typical knowledge worker may not be similar to a business student who holds (or is actively pursuing) higher degrees of education. Additional administration of this experiment outside the academic environment could produce additional insights as well.

Second, the experimental task (searching for journal articles from an academic database) may be suspect. While it is true that searching for academic journal articles is not something most knowledge workers do every day, it is fairly common that most knowledge workers have to search for some form of knowledge objects in their day-to-day activity. The operationalization of this experiment does compare with that common task. As far as the search scenarios are concerned, great efforts were made to ensure the maximum likelihood that a diverse set of scenarios were made available. While the data set was built from a single journal, *MIS Quarterly*, the articles contained within that journal and within the scope of these scenarios ranges widely. Specific effort was made to ensure the scenarios were not too technology based, but rather covered a wide variety of topics. System design was chosen as a representative topic within information systems. The other two topics, user acceptance and risk management, were chosen because they have broader appeal.

Third, in all experimental designs tradeoffs must be made. In this study a repeated measures design was chosen only for the result set size variable. This experiment was specifically designed not to have each subject perform searches with both search interfaces. Others researchers (Speier and Morris, 2003) have chosen repeated measure designs in similar situations. It was a concern that repeating the searches in both search interfaces could confound the results by producing an inflated learning effect. In the Speier and Morris study they did not randomize their task complexity, instead always choosing to perform the low complex task first, and then the more complex task. This was addressed as one of their limitations that could lead to a learning effect. Furthermore, they faced a potential a carryover effect because they chose not to randomly assign their query interfaces. In the case of this experiment, for the one repeated measures variable (result set size), the scenarios were randomized to avoid any carry over, cross over, or learning effect (Girden, 1992).

The lack of an adequate experience measure was an issue for this experiment. There is plenty of theory supporting the idea that experience should play a factor in search tasks (Gregor & Benbasat, 1999; Bedard, 1989), but none of the items used in collecting the data represented a significant measure. This could be a result of a number of different factors. One, there is no agreed upon, validated, scale for measuring experience, thus the researcher was left to develop an experience measurement instrument. Two, even if a good instrument was found for measuring experience, the homogeneity of the subject pool may have prevented this study from detecting a difference. Third, in several of the questions the phrase "in my place of work" was used. This phrase my have been a confounding factor as many of the subjects were full time students. This conjecture is supported by the fact that only one undergraduate subject reported themselves to be an *IS Professional*. Future extension to this study should investigate a better measure for experience.

Future Research

This research just begins to examine the effects of cognitive loading may have on retrieval effectiveness. The discussion of additional future research that follows could help to provide a more complete account of the effects of cognitive loading on retrieval effectiveness. One avenue for future research would be to port the system developed for this experiment to an industry setting. For example, a corporation with a knowledge management initiative might have built a knowledge management system that captures areas of expertise for their employees. It is presumed that a system such as this may have allowed its users to input their own areas of expertise. Furthermore, it is also presumed that the search mechanism implemented in this system is probably keyword based. If this is the case then we have a situation where a keyword search limitation (LaBrie & St. Louis, 2003) could be a problem. One user of the system might put in a skill such as "XML," another as "XML programmer," and another as "extensible markup language." Somebody searching for these people in a text-based keyword search interface would more than likely miss the third individual. The implementation of a visual tree-view hierarchy search interface would possibly reduce this missed opportunity.

Another avenue for future research would be to develop a system that combines the functionality of a keyword and visual search interface. A user might like to initially narrow the amount of information retrieved via one interface and then switch to the other interface for further search refinements. There is no reason this has to be a single search interface option. Technology should allow for the use of both search interfaces interchangeably and simultaneously.

Further extensions to this research could involve the integration of additional feedback during the search process. Several subjects during debriefing sessions made mention that they did not know how well they were doing. They said they might have been more satisfied with the search interfaces if they were given feedback on how well they were achieving their goal. While it may be rare in knowledge management systems to have a predetermined set of correct knowledge objects, if there were such systems then they could possibly benefit from such a feedback mechanism.

Conclusions

The findings from this study suggest that a recognition-based search interface provides superior retrieval accuracy over a recall-based search interface. Theory from cognitive psychology on this aspect of human retrieval skills holds true in a computerized knowledge management environment. The evidence from this simulated document-based KMS shows a 40-50% accuracy gain utilizing a hierarchical visual search interface over a keyword search interface.

Companies that have knowledge management systems with traditional keyword search mechanisms can easily and affordably integrate a visual search interface using existing technology. Most knowledge management systems are based on relational database systems as their underlying storage and retrieval infrastructure. By using a multidimensional database system linked to the KMS relational database system, KMS designers can provide an alternative visual search interface mechanism and instantly reap the rewards of higher accuracy retrieval rates.

Finally the results from this study show that even if you provide a more accurate retrieval mechanism for knowledge objects it does not mean that users will accept it as a viable search solution. Resistance via lower satisfaction, additional time, and/or additional perceived work effort may play a factor in the acceptance of the new search interface. It cannot be ignored that these issues demand further investigation of theories concerning technology acceptance with regards to introduction of a new paradigm for searching within knowledge management systems.

REFERENCES

Alavi, M., & Leidner, D. E. (1999). Knowledge management systems: Issues, challenges, and benefits. *Communications of the Association for Information Systems*, *1(7),* 1-37.

Alavi, M., & Leidner, D. E. (2001). Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, *25(1),* 107-136.

Anderson, J. R. (1995). *Cognitive psychology and its implications* (4^{th} ed.). New York: W. H. Freeman and Company.

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist, 51,* 355-365.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38,* 341-380.

Anderson, J. R., Matessa, M., & Lebiere, C. (1997). ACT-R: A theory of higher level cognition and its relation to visual attention. *Human Computer Interaction*, *12(4),* 439-462.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93(2),* 154-179.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system. In G. H. Bower (Ed.), *The nature of reinforcement* (pp. 66-120). New York: Academic Press.

Baddeley, A. (1994). The magical number seven: Still magical after all these years? *Psychological Review, 101(2),* 353-356.

Ball, M. K. (2002, March). Knowledge management: Intelligence for today's business world. *KMWorld,* S14-S15.

Barki, H., Rivard, S., & Talbot, J. (1988). An information systems keyword classification scheme. *MIS Quarterly, 12(2),* 299-323.

Barki, H., Rivard, S., & Talbot, J. (1993). A keyword classification scheme for IS research literature: An update. *MIS Quarterly, 17(2),* 209-226.

Bedard, J. (1989). Expertise in auditing: Myth or reality? *Accounting Organizations and Society, 14(1/2),* 113-131.

Birkinshaw, J., Nobel, R., & Ridderstrale, J. (2002). Knowledge as a contingency variable: Do the characteristics of knowledge predict organizational structure? *Organizational Science, 13(3),* 274-289.

Blair, D. C. (2002a). The challenge of commercial document retrieval, part I: Major issues, and a framework based on search exhaustivity, determinacy of representation and document collection size. *Information Processing & Management, 38(2),* 273-291.

Blair, D. C. (2002b). The challenge of commercial document retrieval, part II: A strategy for document searching based on identifiable document partitions. *Information Processing & Management, 38(2),* 293-304.

Blair, D. C., & Kimbrough, S. O. (2002). Exemplary documents: A foundation for information retrieval design. *Information Processing & Management, 38(3),* 363-379.

Bower, G. H., & Anderson, J. R. (1973). *Human associative memory.* Washington, DC: Winston & Sons.

Bower, G. H., Clark, J. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorical word lists. *Journal of Verbal Learning & Verbal Behavior, 8,* 323-343.

Brown, J. (1976). An analysis of recognition and recall and of problems in their comparison. In J. Brown (ed.), *Recall and Recognition* (pp. 1-37). New York: John Wiley & Sons.

Buckley, C. (1995). Delphi: A methodology for preferences more than predictions. *Library Management, 16(7),* 6-19.

Business Wire (2003, July 30). Federal knowledge management spending to reach $1.3 billion. Retrieved November 15, 2003 from http://www.lexis-nexis.com.

Chandrasekaran, B., Josephson, J. R., & Benjamins, V. R. (1999). What are ontologies, and why do we need them? *IEEE Intelligent Systems, 14(1),* 20-26.

Chaudhuri, S., & Gravano, L. (1999). Evaluating top-k selection queries. *Proceedings of the VLDB Conference, Edinburgh, Scotland, 25,* 399-410.

Chen, C.-M., & Ling, Y. (2002). A sampling-based estimator for top-k selection queries. *Proceedings of the IEEE International Conference on Data Engineering, 18,* 617 - 627.

Clark, S. E. (1999). Recalling to recognize and recognizing recall. In C. Izawa (ed.), *On Human Memory: Evolution, Progress, and Reflections on the 30th Anniversary of the Atkinson-Shiffrin Model* (pp. 151-164). Mahwah, NJ: Lawrence Erlbaum Associates.

Cochran, W. G. (1968). The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics, 24,* 295-313.

Cochran, W. G. (1983). *Planning and Analysis of Observational Studies.* New York: John Wiley & Sons.

Correa da Silva, F. S., Vasconcelos, W. W., Robertson, D. S., Brilhante, V., de Melo, A. C. V., Finger, M., & Agusti, J. (2002). On the insufficiency of ontologies: problems in knowledge sharing and alternative solutions. *Knowledge-Based Systems, 15(3),* 147-167.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13(3),* 319-340.

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science, 9(3),* 458-467.

Dhaliwal, J. S., & Benbasat, I. (1996). The use and effects of knowledge-based system explanations: Theoretical foundations and a framework for empirical evaluation. *Information Systems Research, 7(3),* 342-361.

Doll, W. J., & Torkzadeh, G. (1988). The measurement of end-user computing satisfaction, *MIS Quarterly, 12(2),* 259–274.

Doll, W. J., Xia, W., & Torkzadeh, G. (1994). Confirmatory factor analysis of the end-user computing satisfaction instrument, *MIS Quarterly, 18(4),* 453-461.

Driscoll, M. P. (2000). *Psychology of learning for instruction,* (2nd ed.). Boston: Allyn and Bacon.

Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology.* (H. A. Ruger & C. E. Bussenues, Trans.). New York: Teachers College, Columbia University. (Original work published 1885)

Edgington, T., Choi, B.-J., Henson, K., Raghu, T. S., & Vinzé, A. S. (in press). Ontology-enabled knowledge management: Adopting ontology to facility knowledge sharing. *Communications of the ACM.*

Fisher, S. L., & Ford, J. K. (1998). Differential effects of learner effort and goal orientation on two learning outcomes, *Personnel Psychology, 51(2),* 397-420.

Freeze, R., Kulkarni, U., & Ravindran, S. (2004). A knowledge management success model: Theoretic development and empirical validation. Manuscript submitted for publication.

Gelderman, M. (2002). Task difficulty, task variability and satisfaction with management support systems, *Information & Management, 39(7),* 593-604.

Gilmour, D. (2003, October). How to fix knowledge management, *Harvard Business Review, 81(10),* 16.

Girdon, E. R. (1992). *ANOVA Repeated Measures.* Newbury Park, CA: Sage Publications.

Gordon, M. D. (1997). It's 10 a.m. do you know where your documents are? The nature and scope of information retrieval problems in business. *Information Processing and Management, 33(1)*, 107–121.

Gordon, M. D., & Moore, S. A. (1999). Depicting the use and purpose of documents to improve information retrieval. *Information Systems Research, 10(1),* 23-37.

Gorla, N., & Walker, G. (1998). Is the lack of keyword synergism inhibiting maturation in the MIS theory? An exploratory study. *Information Processing & Management, 34(2,3),* 325-339.

Gregor, S., & Benbasat, I. (1999). Explanations from intelligent systems: Theoretical foundations and implications for practice. *MIS Quarterly, 23(4),* 497-531.

Grise, M. L., & Gallupe, R .B. (1999). Information overload: Addressing the productivity paradox in face-to-face electronic meetings. *Journal of Management Information Systems, 16(3),* 157-185.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Handcock & N. Meshkati (eds.), *Human mental workload* (pp. 139-183). Amsterdam: North Holland.

Hicks, B. J., Culley, S. J., Allen, R. D., & Mullineux, G. (2002). A framework for the requirements of capturing, storing, & reusing information & knowledge in engineering design. *International Journal of Information Management, 22(4),* 263-280.

Hintze, J. (2001). Number Cruncher Statistical Systems (Version 2004) [Computer software]. Kaysville, UT: WWW.NCSS.COM.

Hoffer, J. A., Prescott, M. B., & McFadden, F. R. (2002). *Modern database management* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Hori, K. (2000). An ontology of strategic knowledge: Key concepts and applications. *Knowledge-Based Systems, 13,* 369-374.

Inmon, W. H. (1996). *Building the data warehouse* (2nd ed.). New York: John Wiley & Sons.

Jordan, P. C. (1986). Effects of an extrinsic reward on intrinsic motivation: A field experiment. *Academy of Management Journal, 29(2),* 405-412.

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Kerlinger, F., & Lee, H. (2000). *Foundations of behavioral research* (4th ed.). Fort Worth, TX: Harcourt College Publishers.

Kimball, R. (1996). *The data warehouse toolkit.* New York: John Wiley & Sons.

Kintsch, W. (1968). Recognition and free recall of organized lists. *Journal of Experimental Psychology, 78(3),* 481-487.

LaBrie, R. C., & St. Louis, R. D. (2003). Barriers of information retrieval from knowledge management systems: An examination of keyword limitations. *Proceedings of the Americas Conference on Information Systems, Tampa, FL, 9,* 2552-2563.

LaBrie, R. C., & St. Louis, R. D. (2004). Intelligent hierarchies for business intelligence information retrieval. Manuscript submitted for publication.

Lambrix, P., & Shahmehri, N. (2000). Querying documents using content, structure and properties. *Journal of Intelligent Information Systems, 15(3),* 287-307.

Lockart, R. S., Craik, F. I. M., & Jacoby, L. (1976). Depth of processing, recognition & recall. In J. Brown (ed.), *Recall and recognition* (pp. 103-130). New York: John Wiley & Sons.

Mao, J.-Y., & Benbasat, I. (2000). The use of explanations in knowledge-based systems: Cognitive perspectives and a process-tracing analysis, *Journal of Management Information Systems, 17(2),* 153-180.

Markus, M. L. (2001). Toward a theory of knowledge reuse: Types of knowledge reuse situations and factors in reuse success. *Journal of Management Information Systems, 18(1),* 57-93.

Marsden, J. R., Pakath, R., & Wibowo, K. (2002). Decision-making under time pressure with different information sources and performance-based financial incentives – Part 2. *Decision Support Systems, 34(1),* 99-124.

Meadow, C.T., & Yuan, W. (1997). Measuring the impact of information: Defining the concepts. *Information Processing & Management, 33(6),* 697-715.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limitations on our capacity for processing information. *Psychological Review, 63,* 81-97.

Moore, J. C., Richmond, W. B., & Whinston, A. B. (1990). A decision-theoretic approach to information retrieval. *ACM Transactions on Database Systems, 15(3),* 311-340.

Moore, R. W. (2002, April). The preservation of data, information, & knowledge. *Proceedings of the World Library Summit, Singapore.*

Morris, M. G., Speier, C., & Hoffer, J. A. (1999). An examination of procedural and object-oriented systems analysis methods: Does prior experience help or hinder performance? *Decision Sciences, 30(1),* 107-137.

Neufeld, D. J., & Staples, S. (2002). An investigation of information systems topic coverage in IS versus functional area business journals. *Americas Conference on Information Systems, 8, Dallas, TX,* 1759-1760.

Parikh, M., Fazlollahi, B., & Verma, S. (2001). The effectiveness of decisional guidance: An empirical evaluation. *Decision Sciences, 32(2),* 303-331.

Raisinghani, M. S. (2000). Knowledge management: A cognitive perspective on business and education. *Journal of Knowledge Management, 18(2),* 105-112.

Schultze, U., & Leidner, D. E. (2002). Studying knowledge management in information systems research: Discourses and theoretical assumptions. *MIS Quarterly, 26(3),* 213-242.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Simon, H. A. (1962). The architecture of complexity: Hierarchical systems. *Proceedings of the Philosophical Society, 106,* 467-482.

Simon, H. A. (1976). How big is a chunk? *Science, 183,* 482-488.

Simon, H. A. (2001). *The sciences of the artificial* (3rd ed.). Cambridge, MA: The MIT Press.

Smith, V. L. (1976). Experimental economics: Induced value theory. *The American Economic Review, 66(2),* 274-279.

Speier, C., & Morris, M. G. (2003). The influence of query interface design on decision-making performance, *MIS Quarterly, 27(3),* 397-423.

Spiegler, I. (2000). Knowledge management: A new idea or a recycled concept? *Communications of the Association for Information Systems, 3,* 1-23.

Sprague, R. H. (1995). Electronic document management: Challenges and opportunities for information systems managers. *MIS Quarterly, 19(1),* 29-49.

Swanson, E. B., & Ramiller, N. C. (1993). Information systems research thematics: Submissions to a new journal, 1987-1992. *Information Systems Research, 4(4),* 299-330.

Swartout, W., & Tate, A. (1999). Ontologies. *IEEE Intelligent Systems, 14(1),* 18-19.

Todd, P., & Benbazat, I. (1992). The use of information in decision making: An experimental investigation of the impact of computer-based decision aids. *MIS Quarterly, 16(3),* 373-393.

Teskey, F. N. (1989). User models and world models for data, information, and knowledge. *Information Processing & Management, 25(1),* 7-14.

Trochim, W. M. K. (2001). *The research methods knowledge base* (2nd ed.). Cincinnati, OH: Atomic Dog Publishing.

Van Beveren, J. (2002). A model of knowledge acquisition that refocuses knowledge management. *Journal of Knowledge Management, 6(1),* 18-22.

Vessey, I. (1994). The effect of information presentation on decision making: A cost-benefit analysis. *Information & Management, 27(2),* 103-119.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus table literature. *Decision Sciences, 22(2),* 219-240.

Vessey, I., Ramesh, V., & Glass, R. L. (2002). Research in information systems: An empirical study of diversity in the discipline and its journals. *Journal of Management Information Systems, 19(2),* 129-174.

Wand, Y., Storey, V. C., & Weber, R. (2000). An ontological analysis of the relationship construct in conceptual modeling. *ACM Transactions on Database System, 24(4),* 494-528.

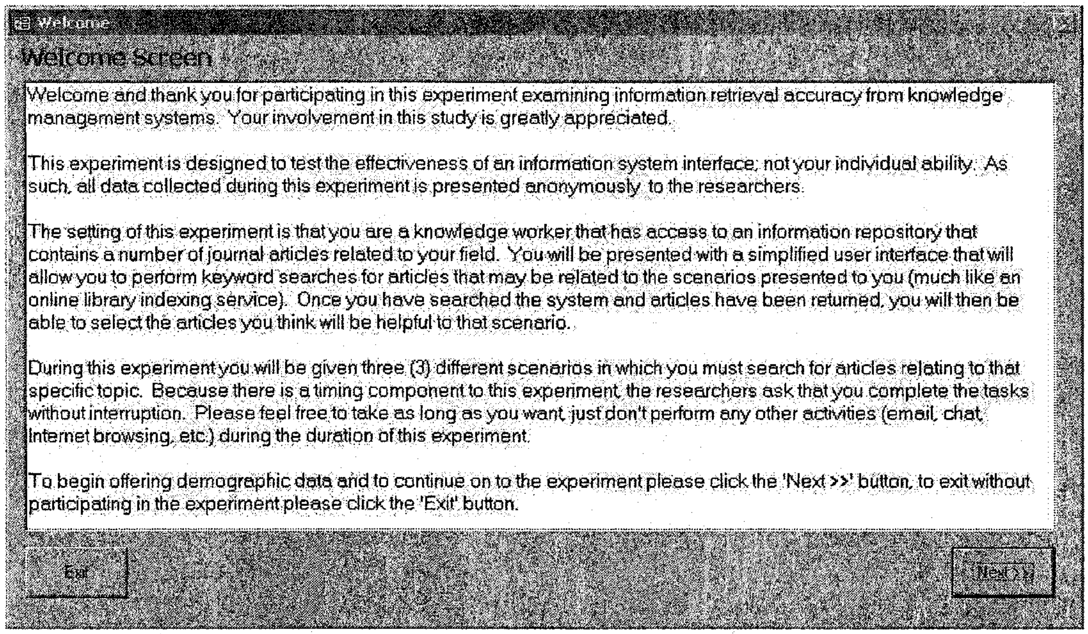
Wand, Y., & Weber, R. (1990). An ontological model of an information system. *IEEE Transactions on Software Engineering, 16(11),* 1282-1293.

Weber, R. (2003). Theoretically speaking. *MIS Quarterly, 27(3),* iii-xii.

Wierwille, W. W., & Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors, 35(2),* 263-281.

APPENDIX A

KEYWORD SEARCH SYSTEM



Welcome

Welcome Screen

Welcome and thank you for participating in this experiment examining information retrieval accuracy from knowledge management systems. Your involvement in this study is greatly appreciated.

This experiment is designed to test the effectiveness of an information system interface, not your individual ability. As such, all data collected during this experiment is presented anonymously to the researchers.

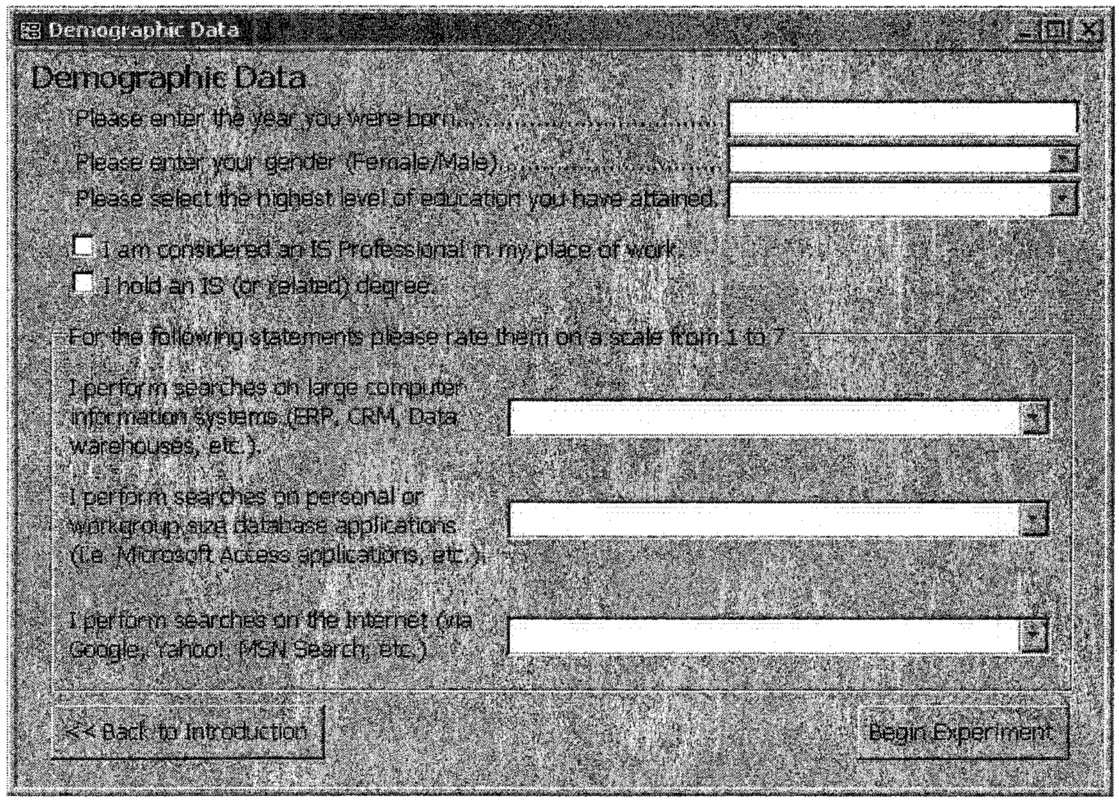
The setting of this experiment is that you are a knowledge worker that has access to an information repository that contains a number of journal articles related to your field. You will be presented with a simplified user interface that will allow you to perform keyword searches for articles that may be related to the scenarios presented to you (much like an online library indexing service). Once you have searched the system and articles have been returned, you will then be able to select the articles you think will be helpful to that scenario.

During this experiment you will be given three (3) different scenarios in which you must search for articles relating to that specific topic. Because there is a timing component to this experiment the researchers ask that you complete the tasks without interruption. Please feel free to take as long as you want, just don't perform any other activities (email, chat, Internet browsing, etc.) during the duration of this experiment.

To begin offering demographic data and to continue on to the experiment please click the 'Next >>' button, to exit without participating in the experiment please click the 'Exit' button.

Exit

Next >>



Demographic Data

Demographic Data

Please enter the year you were born...

Please enter your gender (Female/Male)...

Please select the highest level of education you have attained.

I am considered an IS Professional in my place of work.

I hold an IS (or related) degree.

For the following statements please rate them on a scale from 1 to 7

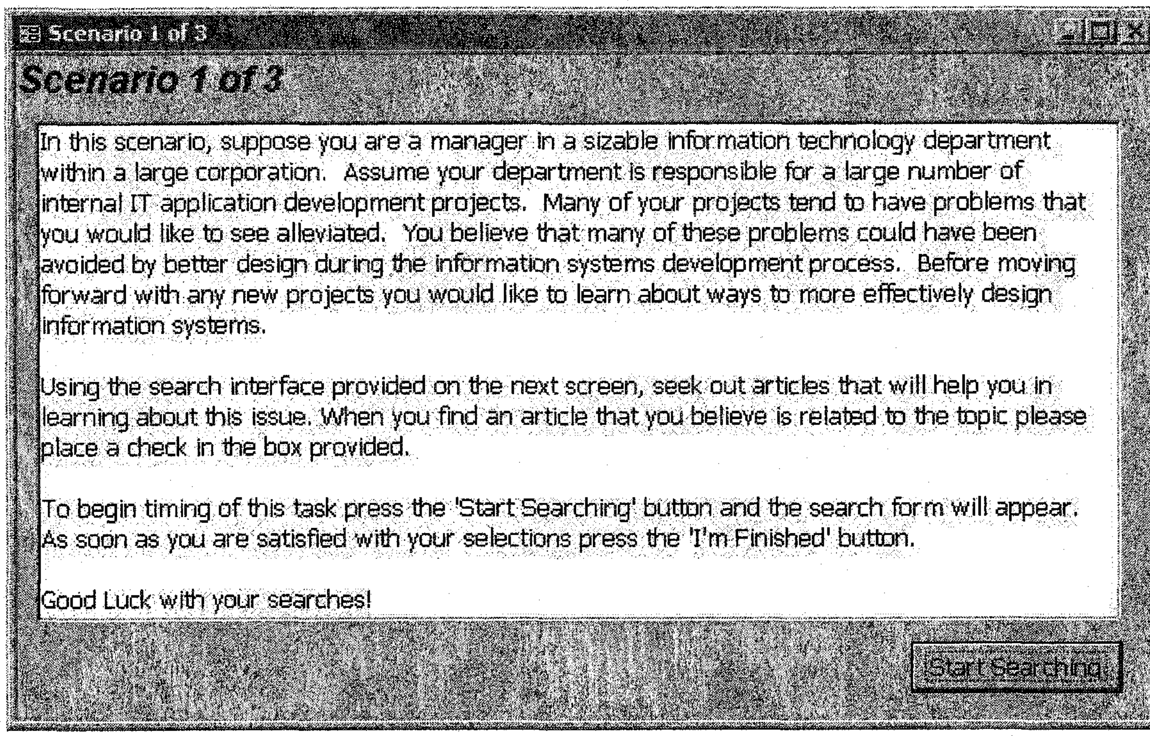
I perform searches on large computer information systems (ERP, CRM, Data warehouses, etc.).

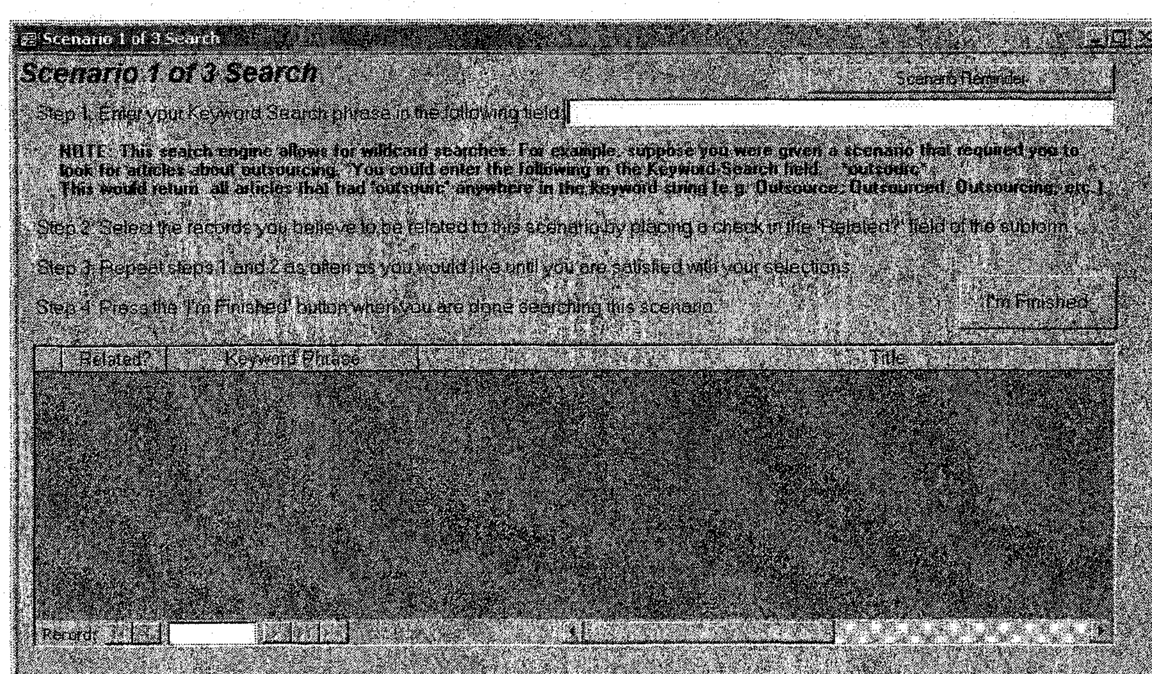
I perform searches on personal or workgroup size database applications (i.e. Microsoft Access applications, etc.).

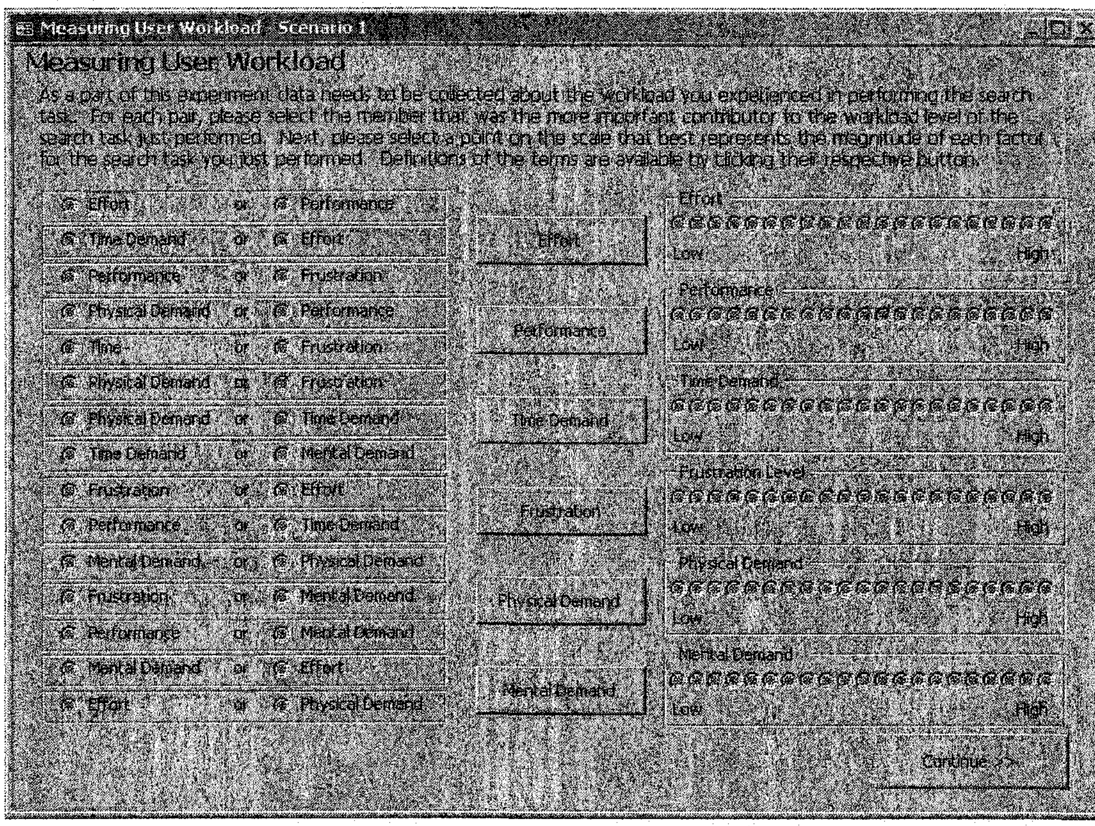
I perform searches on the Internet (via Google, Yahoo!, MSN Search, etc.)

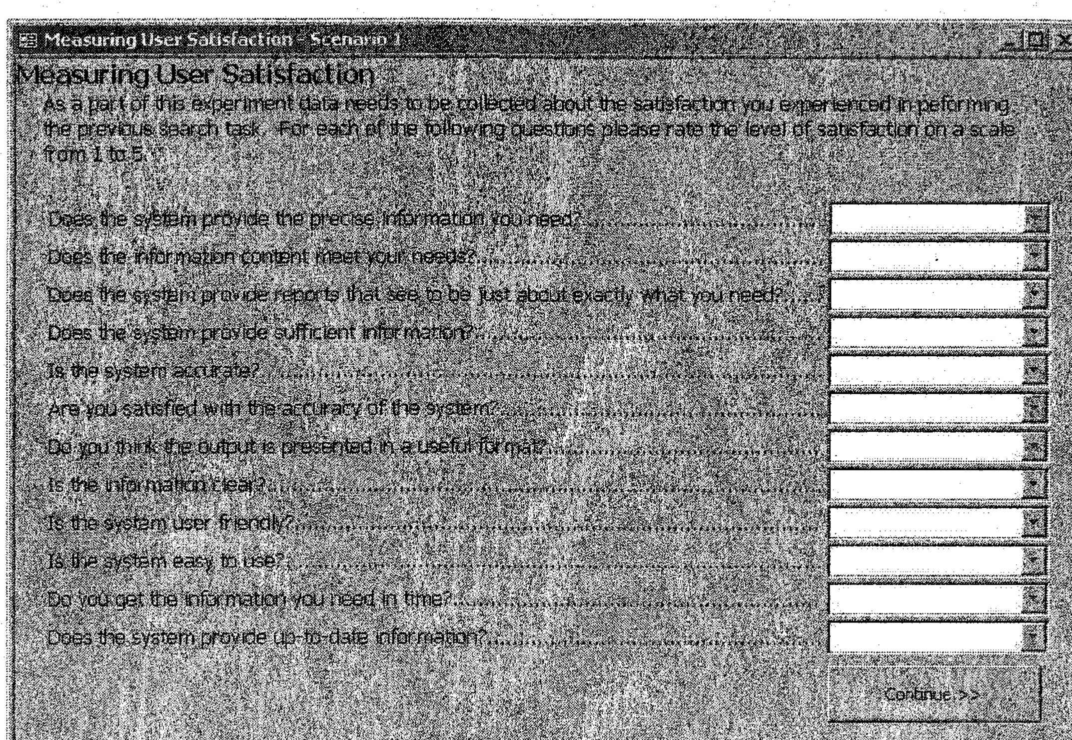
<< Back to Introduction

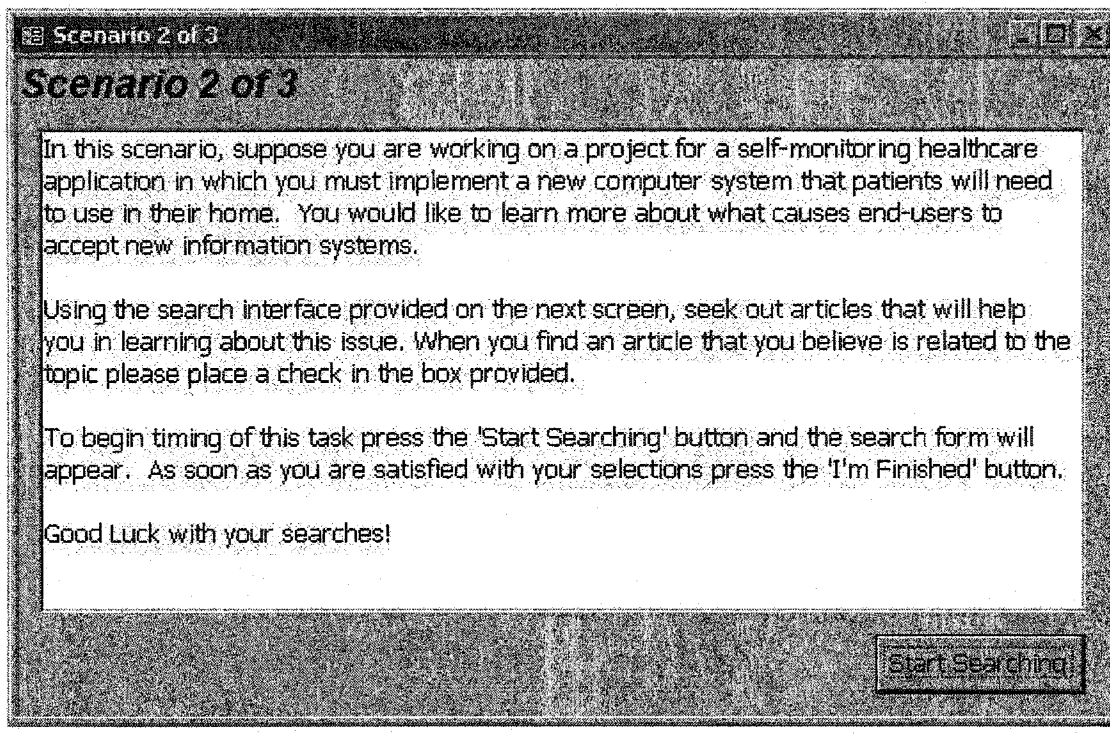
Begin Experiment

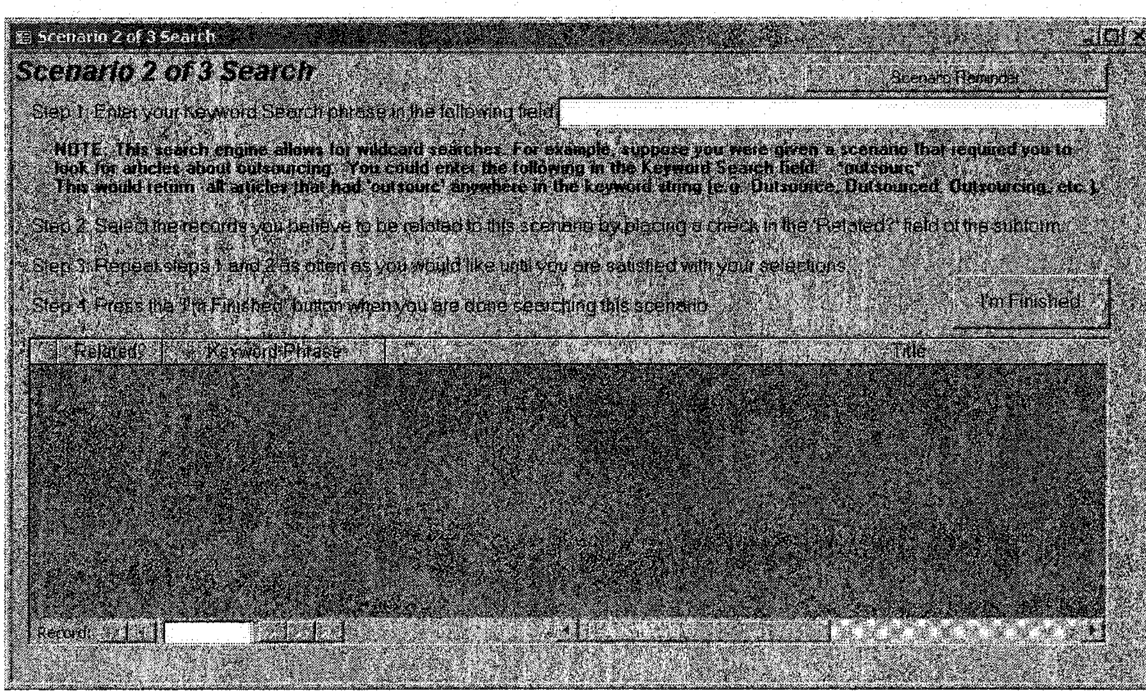

Scenario 1 of 3
Scenario 1 of 3
In this scenario, suppose you are a manager in a sizable information technology department within a large corporation. Assume your department is responsible for a large number of internal IT application development projects. Many of your projects tend to have problems that you would like to see alleviated. You believe that many of these problems could have been avoided by better design during the information systems development process. Before moving forward with any new projects you would like to learn about ways to more effectively design information systems.
Using the search interface provided on the next screen, seek out articles that will help you in learning about this issue. When you find an article that you believe is related to the topic please place a check in the box provided.
To begin timing of this task press the 'Start Searching' button and the search form will appear. As soon as you are satisfied with your selections press the 'I'm Finished' button.
Good Luck with your searches!
Start Searching

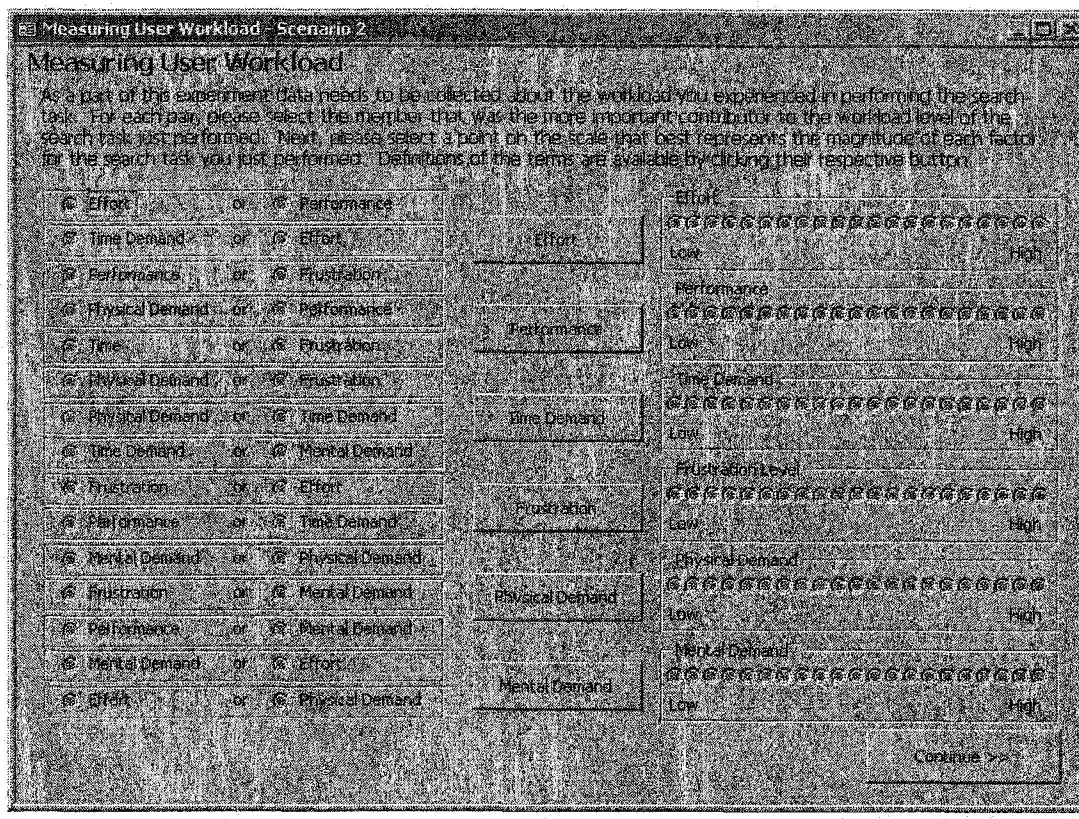

Scenario 1 of 3 Search
Scenario 1 of 3 Search
Scenario Reminder
Step 1: Enter your Keyword Search phrase in the following field:
Step 2: Select the records you believe to be related to this scenario by placing a check in the 'Related?' field of the subform.
Step 3: Repeat steps 1 and 2 as often as you would like until you are satisfied with your selections.
Step 4: Press the 'I'm Finished' button when you are done searching this scenario.
I'm Finished
Related?
Keyword Phrase
Title
Record:

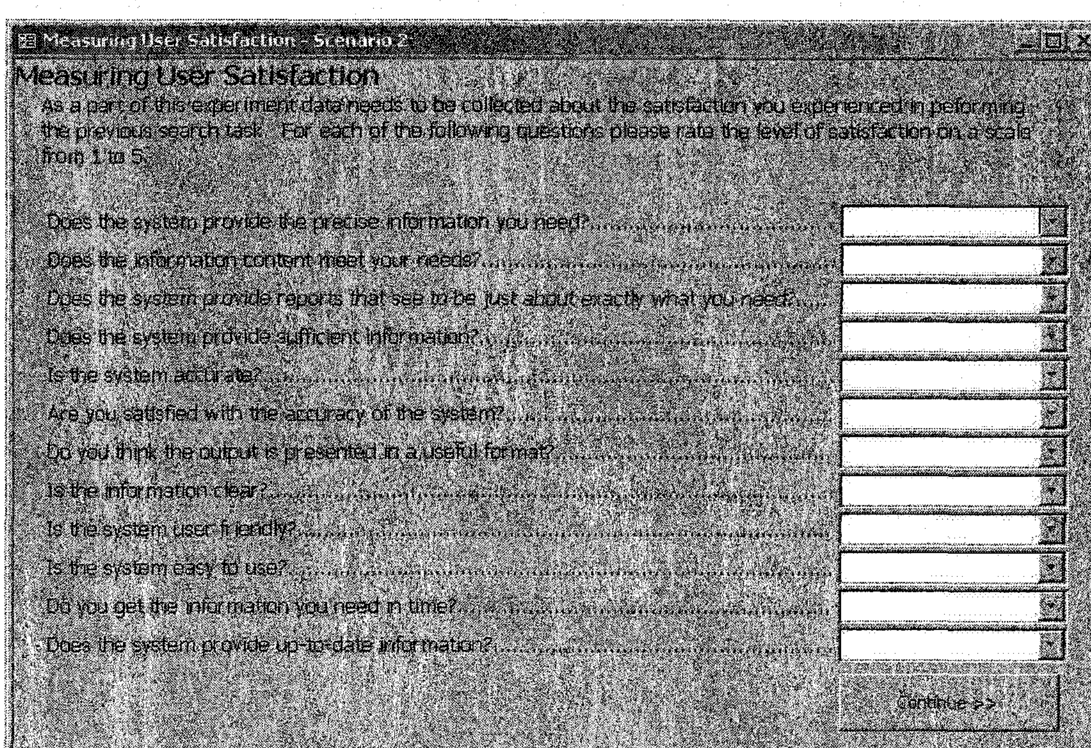

Measuring User Workload - Scenario 1
Measuring User Workload
As a part of this experiment data needs to be collected about the workload you experienced in performing the search task. For each pair, please select the member that was the more important contributor to the workload level of the search task just performed. Next, please select a point on the scale that best represents the magnitude of each factor for the search task you just performed. Definitions of the terms are available by clicking their respective button.
Effort or Performance
Time Demand or Effort
Performance or Frustration
Physical Demand or Performance
Time or Frustration
Physical Demand or Frustration
Physical Demand or Time Demand
Time Demand or Mental Demand
Frustration or Effort
Performance or Time Demand
Mental Demand or Physical Demand
Frustration or Mental Demand
Performance or Mental Demand
Mental Demand or Effort
Effort or Physical Demand
Effort
Performance
Time Demand
Frustration
Physical Demand
Mental Demand
Effort
Low
High
Performance
Low
High
Time Demand
Low
High
Frustration Level
Low
High
Physical Demand
Low
High
Mental Demand
Low
High
Continue >>

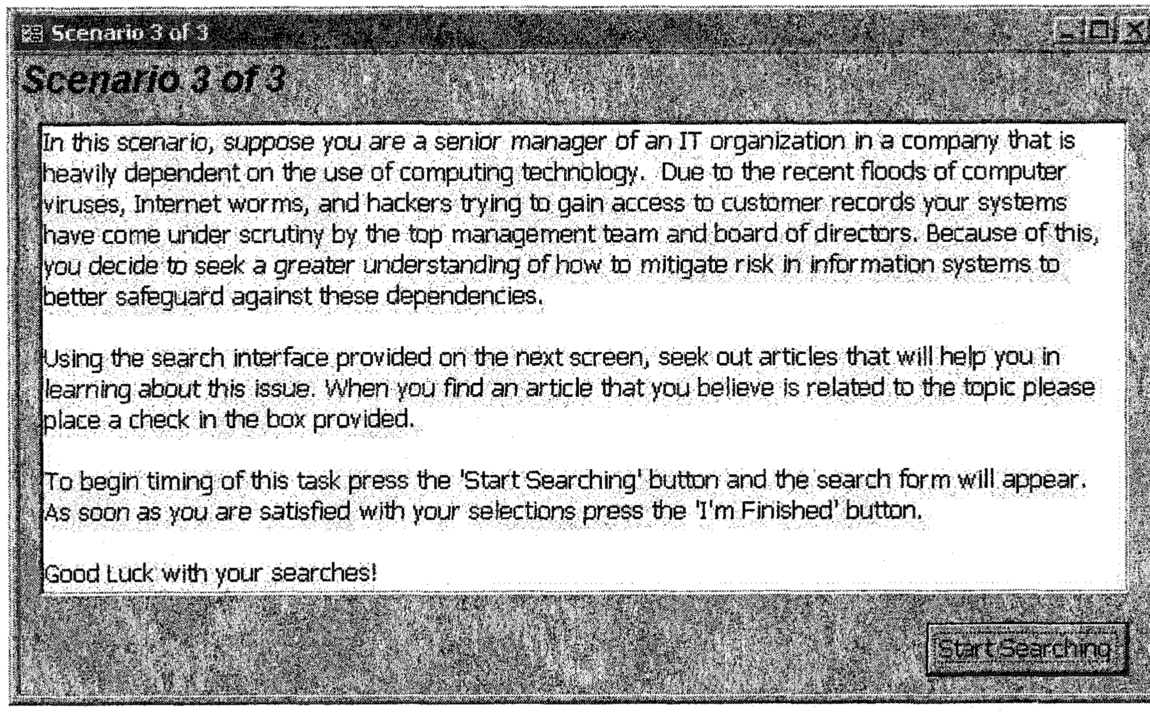

Measuring User Satisfaction - Scenario 1
Measuring User Satisfaction
As a part of this experiment data needs to be collected about the satisfaction you experienced in peforming the previous search task. For each of the following questions please rate the level of satisfaction on a scale from 1 to 5.
Does the system provide the precise information you need?
Does the information content meet your needs?
Does the system provide reports that see to be just about exactly what you need?
Does the system provide sufficient information?
Is the system accurate?
Are you satisfied with the accuracy of the system?
Do you think the output is presented in a useful format?
Is the information clear?
Is the system user friendly?
Is the system easy to use?
Do you get the information you need in time?
Does the system provide up-to-date information?
Continue >>

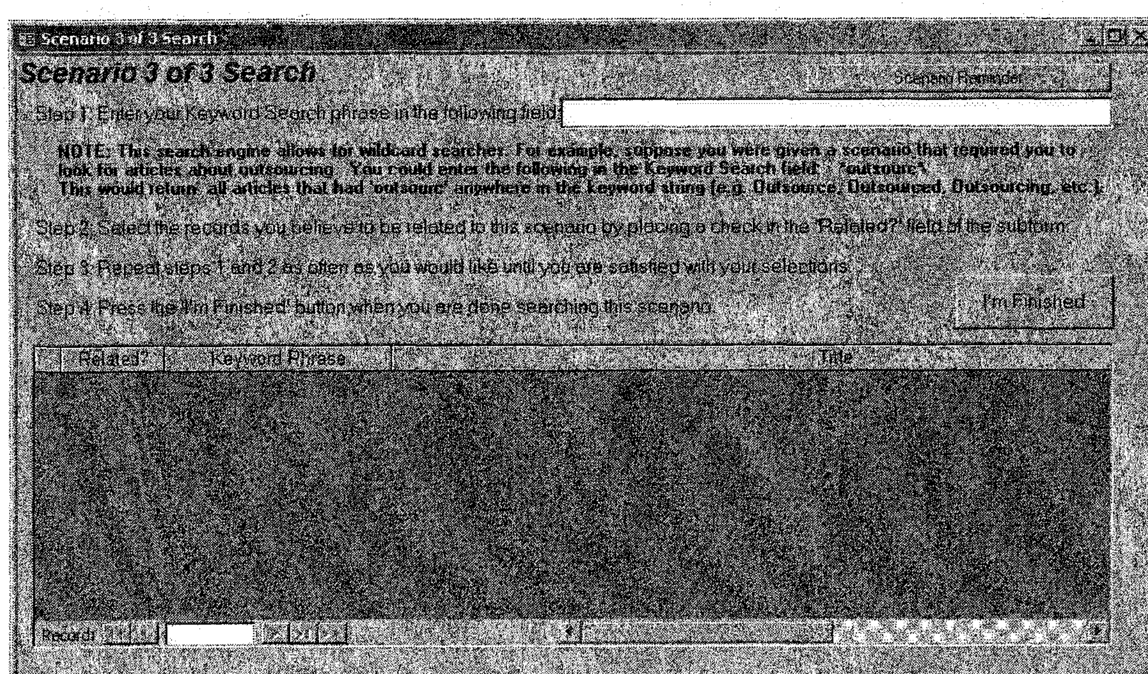

Scenario 2 of 3
Scenario 2 of 3
In this scenario, suppose you are working on a project for a self-monitoring healthcare application in which you must implement a new computer system that patients will need to use in their home. You would like to learn more about what causes end-users to accept new information systems.
Using the search interface provided on the next screen, seek out articles that will help you in learning about this issue. When you find an article that you believe is related to the topic please place a check in the box provided.
To begin timing of this task press the 'Start Searching' button and the search form will appear. As soon as you are satisfied with your selections press the 'I'm Finished' button.
Good Luck with your searches!
Start Searching

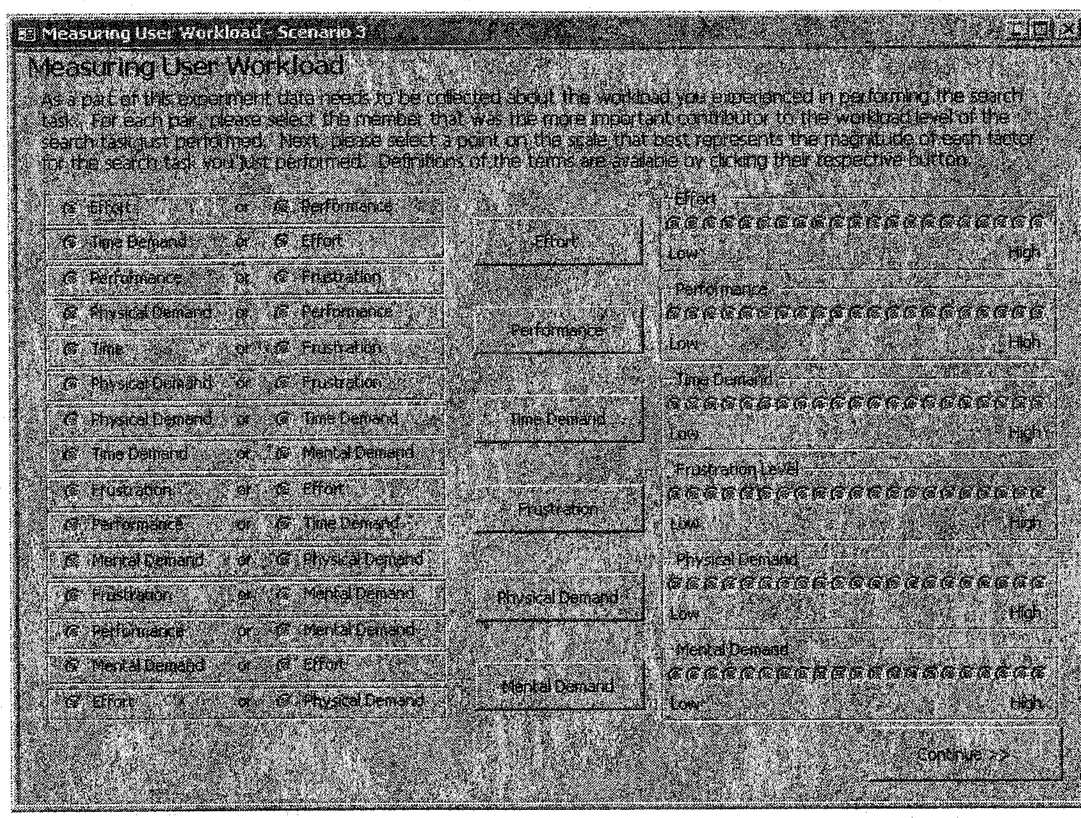

Scenario 2 of 3 Search
Scenario 2 of 3 Search
Scenario Reminder
Step 1: Enter your Keyword Search phrase in the following field:
NOTE: This search engine allows for wildcard searches. For example, suppose you were given a scenario that required you to look for articles about outsourcing. You could enter the following in the Keyword Search field: *outsourc*
This would return all articles that had 'outsourc' anywhere in the keyword string (e.g. Outsource, Outsourced, Outsourcing, etc.).
Step 2: Select the records you believe to be related to this scenario by placing a check in the 'Related?' field of the subform.
Step 3: Repeat steps 1 and 2 as often as you would like until you are satisfied with your selections.
Step 4: Press the 'I'm Finished' button when you are done searching this scenario.
I'm Finished
Related?
Keyword Phrase
Title
Record:

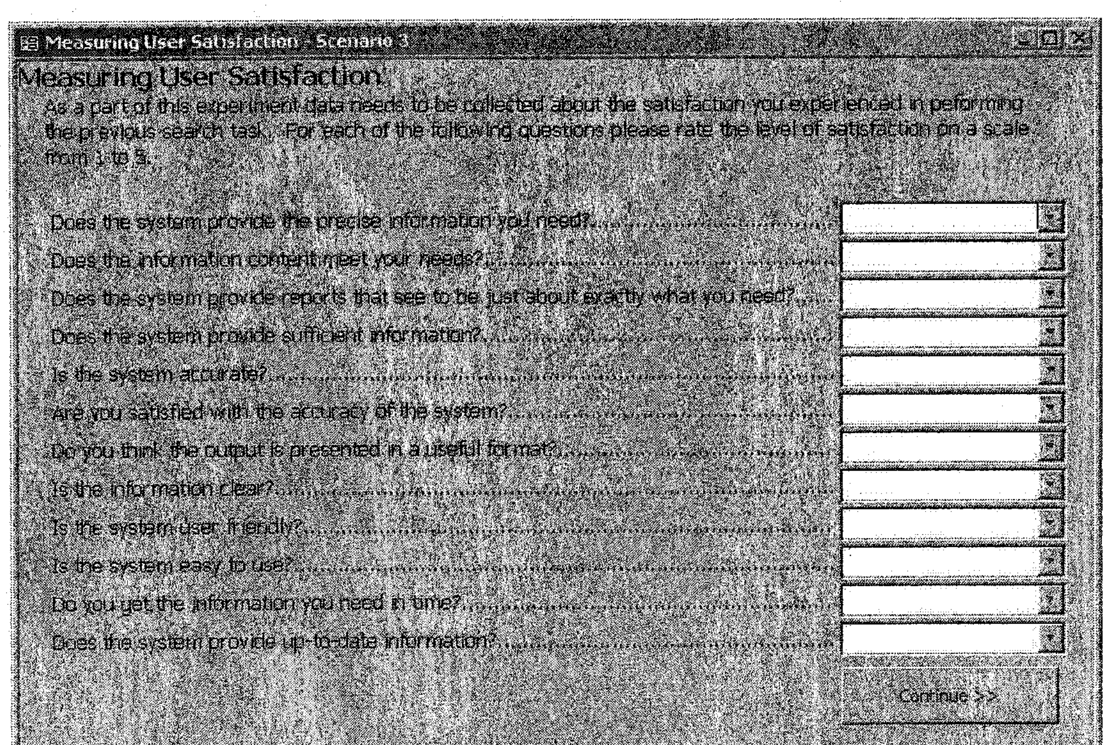

Measuring User Workload - Scenario 2
Measuring User Workload
As a part of this experiment data needs to be collected about the workload you experienced in performing the search task. For each pair, please select the member that was the more important contributor to the workload level of the search task just performed. Next, please select a point on the scale that best represents the magnitude of each factor for the search task you just performed. Definitions of the terms are available by clicking their respective button.
Effort or Performance
Time Demand or Effort
Performance or Frustration
Physical Demand or Performance
Time or Frustration
Physical Demand or Frustration
Physical Demand or Time Demand
Time Demand or Mental Demand
Frustration or Effort
Performance or Time Demand
Mental Demand or Physical Demand
Frustration or Mental Demand
Performance or Mental Demand
Mental Demand or Effort
Effort or Physical Demand
Effort
Performance
Time Demand
Frustration
Physical Demand
Mental Demand
Effort
Low High
Performance
Low High
Time Demand
Low High
Frustration Level
Low High
Physical Demand
Low High
Mental Demand
Low High
Continue >>

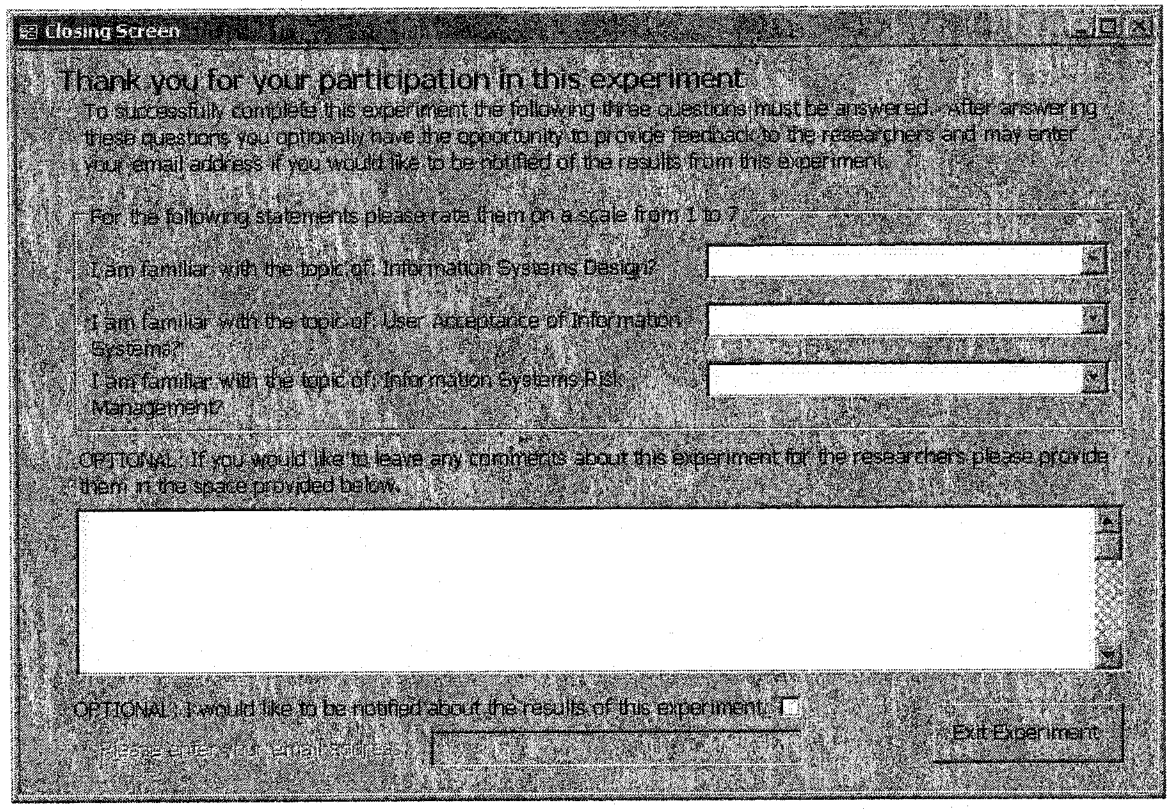

Measuring User Satisfaction - Scenario 2
Measuring User Satisfaction
As a part of this experiment data needs to be collected about the satisfaction you experienced in peforming the previous search task. For each of the following questions please rate the level of satisfaction on a scale from 1 to 5.
Does the system provide the precise information you need?
Does the information content meet your needs?
Does the system provide reports that see to be just about exactly what you need?
Does the system provide sufficient information?
Is the system accurate?
Are you satisfied with the accuracy of the system?
Do you think the output is presented in a useful format?
Is the information clear?
Is the system user friendly?
Is the system easy to use?
Do you get the information you need in time?
Does the system provide up-to-date information?
Continue >>


Scenario 3 of 3
Scenario 3 of 3
In this scenario, suppose you are a senior manager of an IT organization in a company that is heavily dependent on the use of computing technology. Due to the recent floods of computer viruses, Internet worms, and hackers trying to gain access to customer records your systems have come under scrutiny by the top management team and board of directors. Because of this, you decide to seek a greater understanding of how to mitigate risk in information systems to better safeguard against these dependencies.
Using the search interface provided on the next screen, seek out articles that will help you in learning about this issue. When you find an article that you believe is related to the topic please place a check in the box provided.
To begin timing of this task press the 'Start Searching' button and the search form will appear. As soon as you are satisfied with your selections press the 'I'm Finished' button.
Good Luck with your searches!
Start Searching


Scenario 3 of 3 Search
Scenario 3 of 3 Search
Scenario Reminder
Step 1: Enter your Keyword Search phrase in the following field:
NOTE: This search engine allows for wildcard searches. For example, suppose you were given a scenario that required you to look for articles about outsourcing. You could enter the following in the Keyword Search field: *outsourc*
This would return all articles that had 'outsourc' anywhere in the keyword string (e.g. Outsource, Outsourced, Outsourcing, etc.).
Step 2: Select the records you believe to be related to this scenario by placing a check in the 'Related?' field of the subform.
Step 3: Repeat steps 1 and 2 as often as you would like until you are satisfied with your selections.
Step 4: Press the 'I'm Finished' button when you are done searching this scenario.
I'm Finished
Related?
Keyword Phrase
Title
Record:

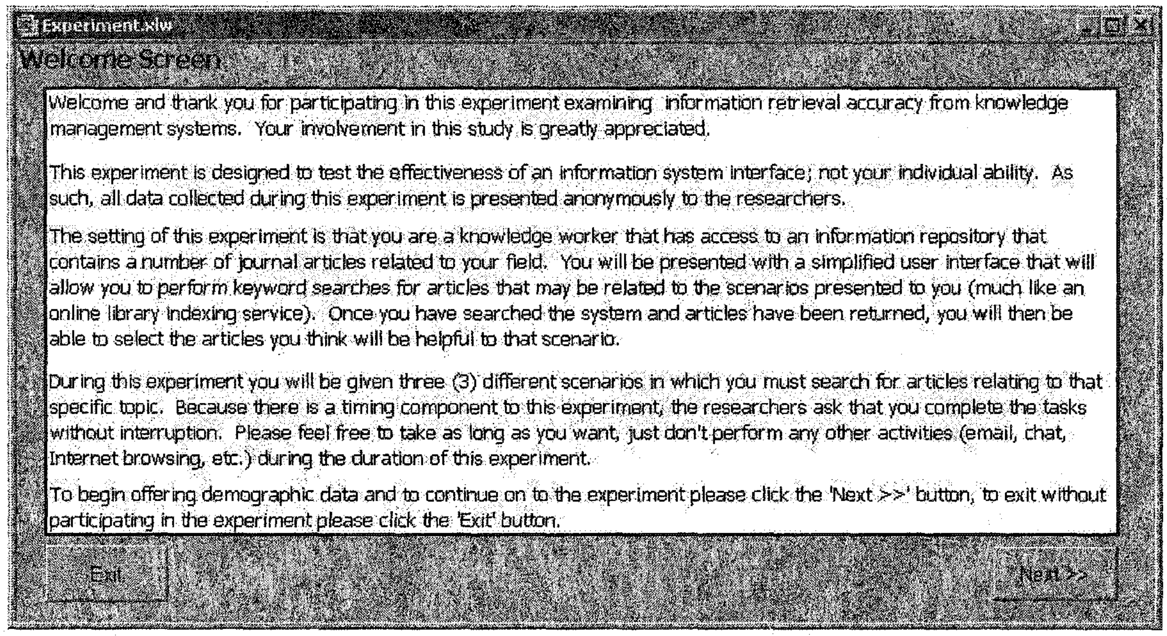

Measuring User Workload - Scenario 3
Measuring User Workload
As a part of this experiment data needs to be collected about the workload you experienced in performing the search task. For each pair, please select the member that was the more important contributor to the workload level of the search task just performed. Next, please select a point on the scale that best represents the magnitude of each factor for the search task you just performed. Definitions of the terms are available by clicking their respective button.
Effort or Performance
Time Demand or Effort
Performance or Frustration
Physical Demand or Performance
Time or Frustration
Physical Demand or Frustration
Physical Demand or Time Demand
Time Demand or Mental Demand
Frustration or Effort
Performance or Time Demand
Mental Demand or Physical Demand
Frustration or Mental Demand
Performance or Mental Demand
Mental Demand or Effort
Effort or Physical Demand
Effort
Performance
Time Demand
Frustration
Physical Demand
Mental Demand
Effort
Performance
Time Demand
Frustration Level
Physical Demand
Mental Demand
Low
High
Continue >>

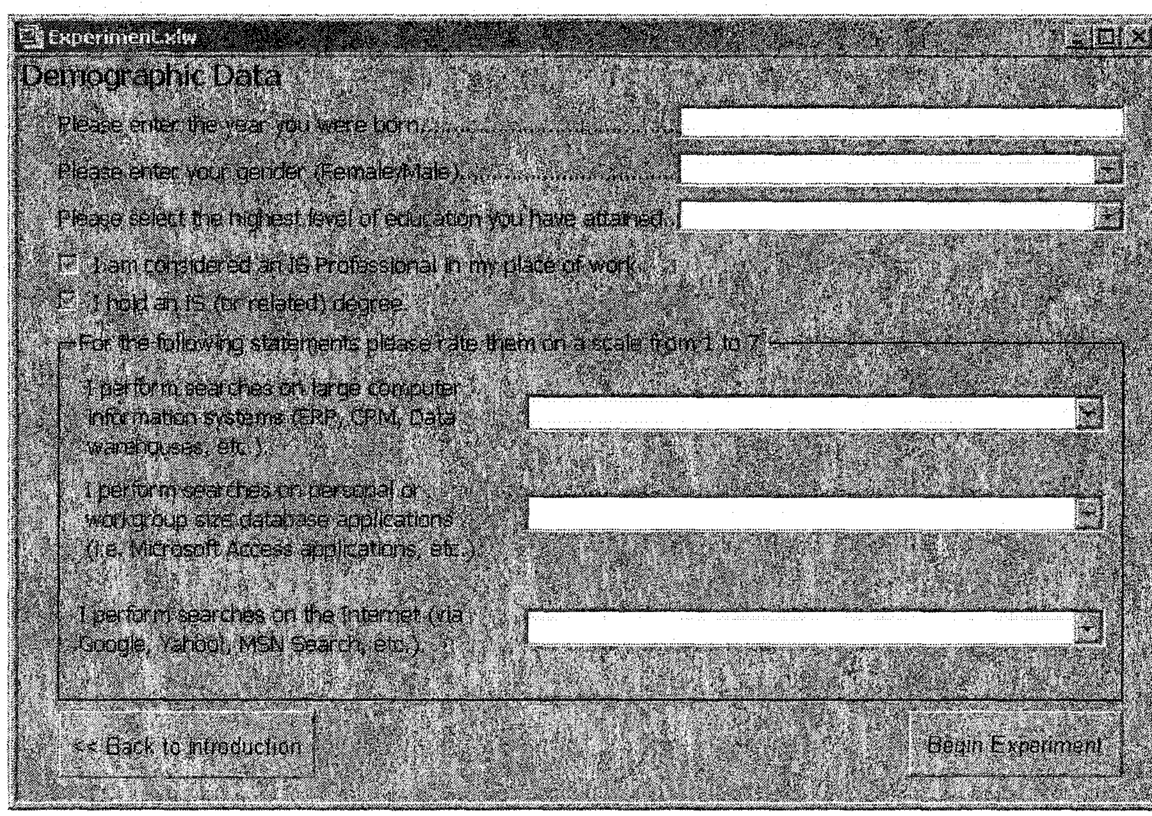

Measuring User Satisfaction - Scenario 3
Measuring User Satisfaction
As a part of this experiment data needs to be collected about the satisfaction you experienced in peforming the previous search task. For each of the following questions please rate the level of satisfaction on a scale from 1 to 5.
Does the system provide the precise information you need?
Does the information content meet your needs?
Does the system provide reports that see to be just about exactly what you need?
Does the system provide sufficient information?
Is the system accurate?
Are you satisfied with the accuracy of the system?
Do you think the output is presented in a useful format?
Is the information clear?
Is the system user friendly?
Is the system easy to use?
Do you get the information you need in time?
Does the system provide up-to-date information?
Continue >>

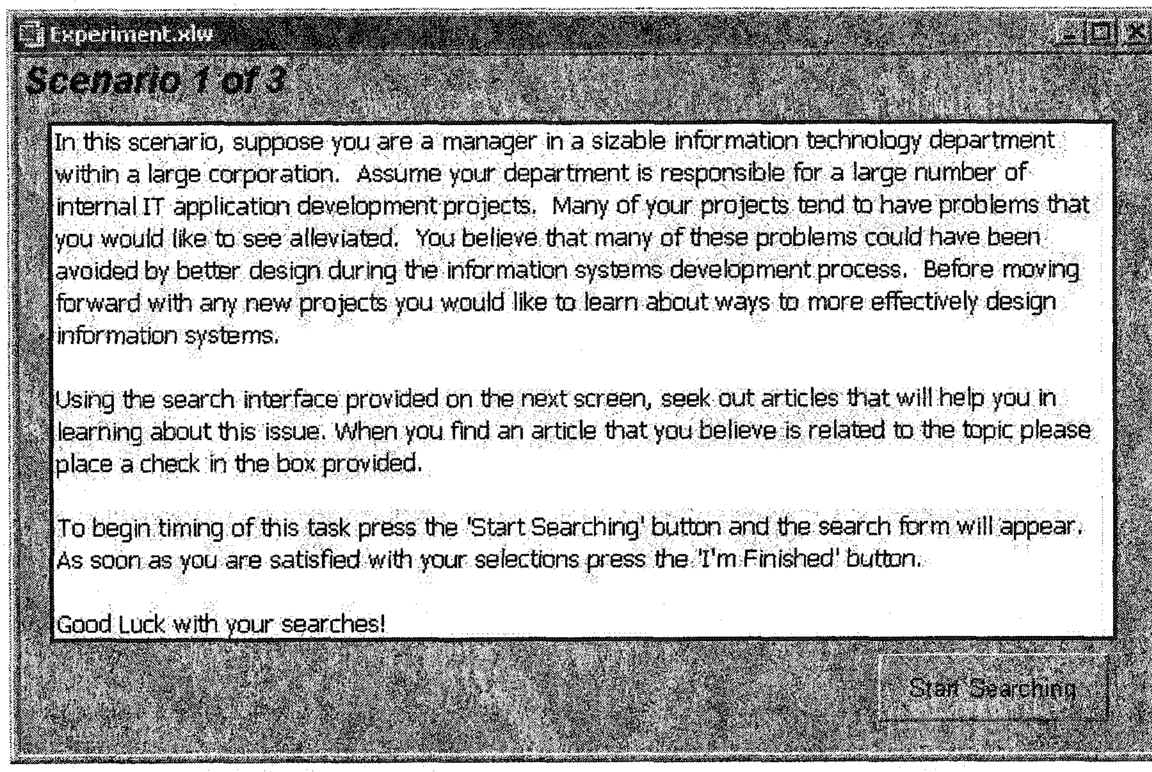

Closing Screen
Thank you for your participation in this experiment
To successfully complete this experiment the following three questions must be answered. After answering these questions you optionally have the opportunity to provide feedback to the researchers and may enter your email address if you would like to be notified of the results from this experiment.
For the following statements please rate them on a scale from 1 to 7
I am familiar with the topic of: Information Systems Design?
I am familiar with the topic of: User Acceptance of Information Systems?*
I am familiar with the topic of: Information Systems Risk Management?
OPTIONAL: If you would like to leave any comments about this experiment for the researchers please provide them in the space provided below.
OPTIONAL: I would like to be notified about the results of this experiment.
Exit Experiment

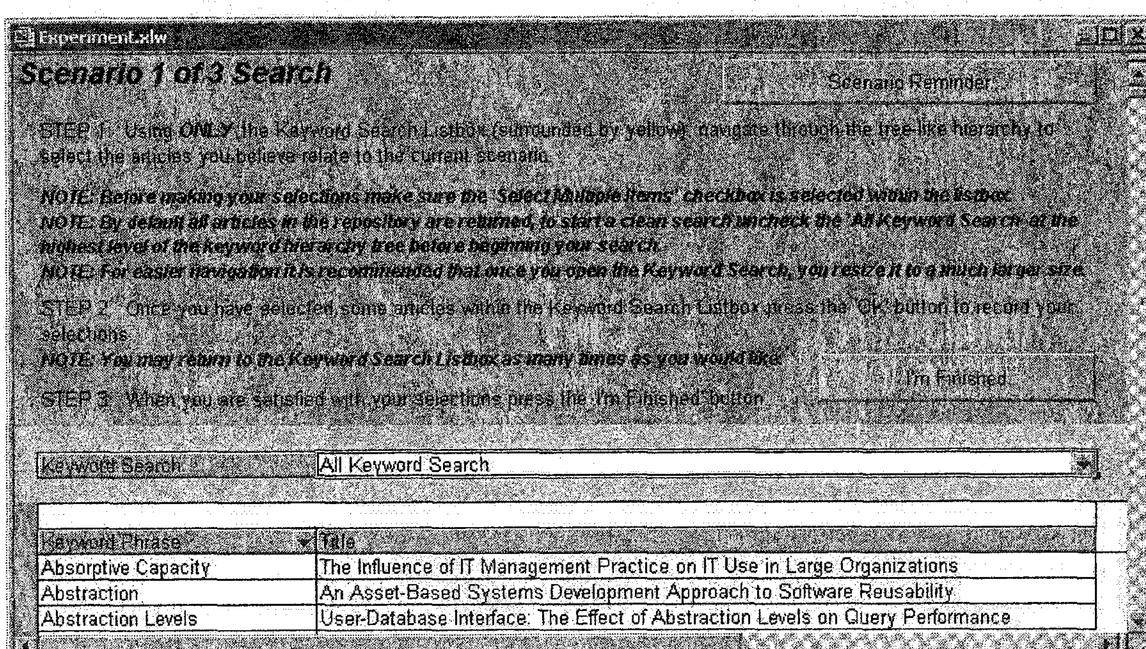
APPENDIX B

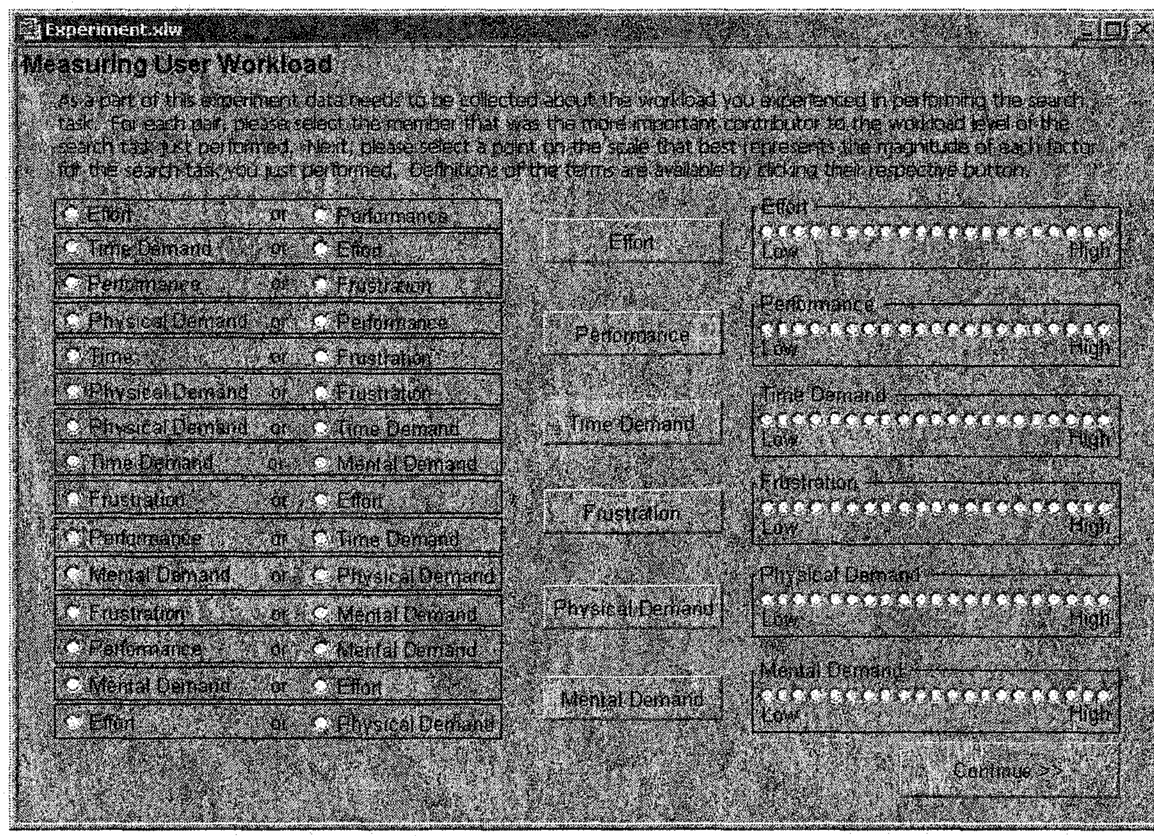
VISUAL SEARCH SYSTEM

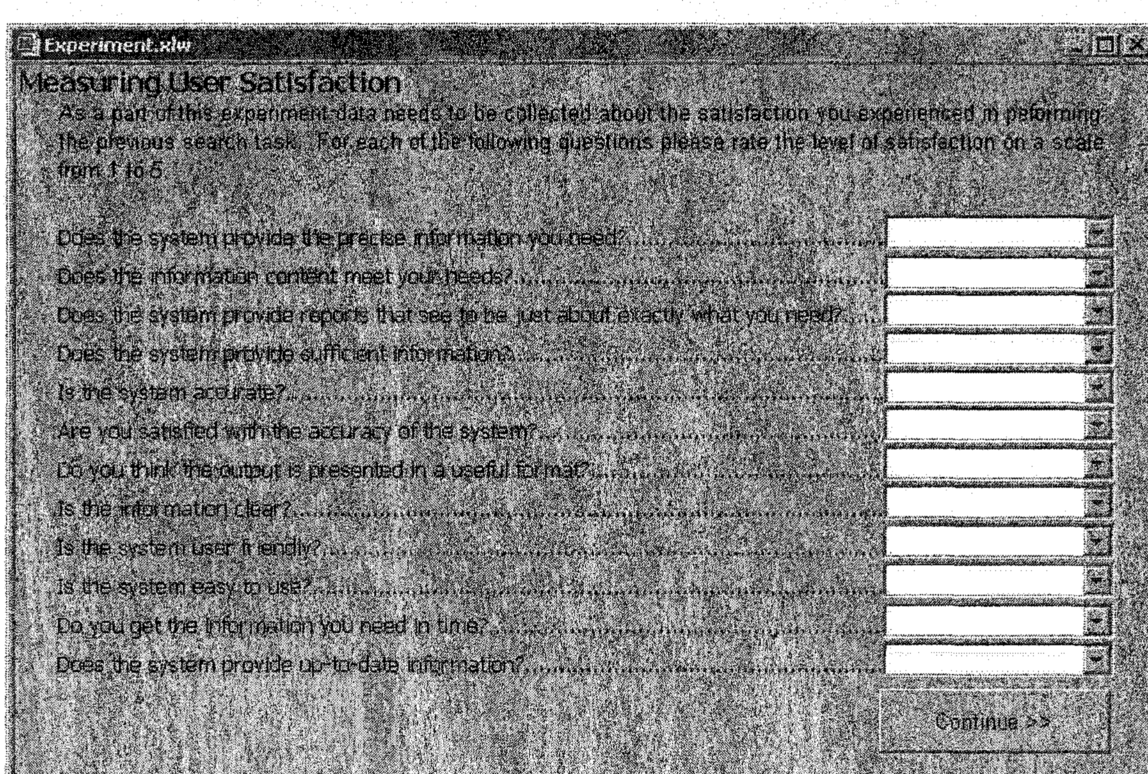

Experiment.xlw
Welcome Screen
Welcome and thank you for participating in this experiment examining information retrieval accuracy from knowledge management systems. Your involvement in this study is greatly appreciated.
This experiment is designed to test the effectiveness of an information system interface; not your individual ability. As such, all data collected during this experiment is presented anonymously to the researchers.
The setting of this experiment is that you are a knowledge worker that has access to an information repository that contains a number of journal articles related to your field. You will be presented with a simplified user interface that will allow you to perform keyword searches for articles that may be related to the scenarios presented to you (much like an online library indexing service). Once you have searched the system and articles have been returned, you will then be able to select the articles you think will be helpful to that scenario.
During this experiment you will be given three (3) different scenarios in which you must search for articles relating to that specific topic. Because there is a timing component to this experiment, the researchers ask that you complete the tasks without interruption. Please feel free to take as long as you want, just don't perform any other activities (email, chat, Internet browsing, etc.) during the duration of this experiment.
To begin offering demographic data and to continue on to the experiment please click the 'Next >>' button, to exit without participating in the experiment please click the 'Exit' button.
Exit
Next >>

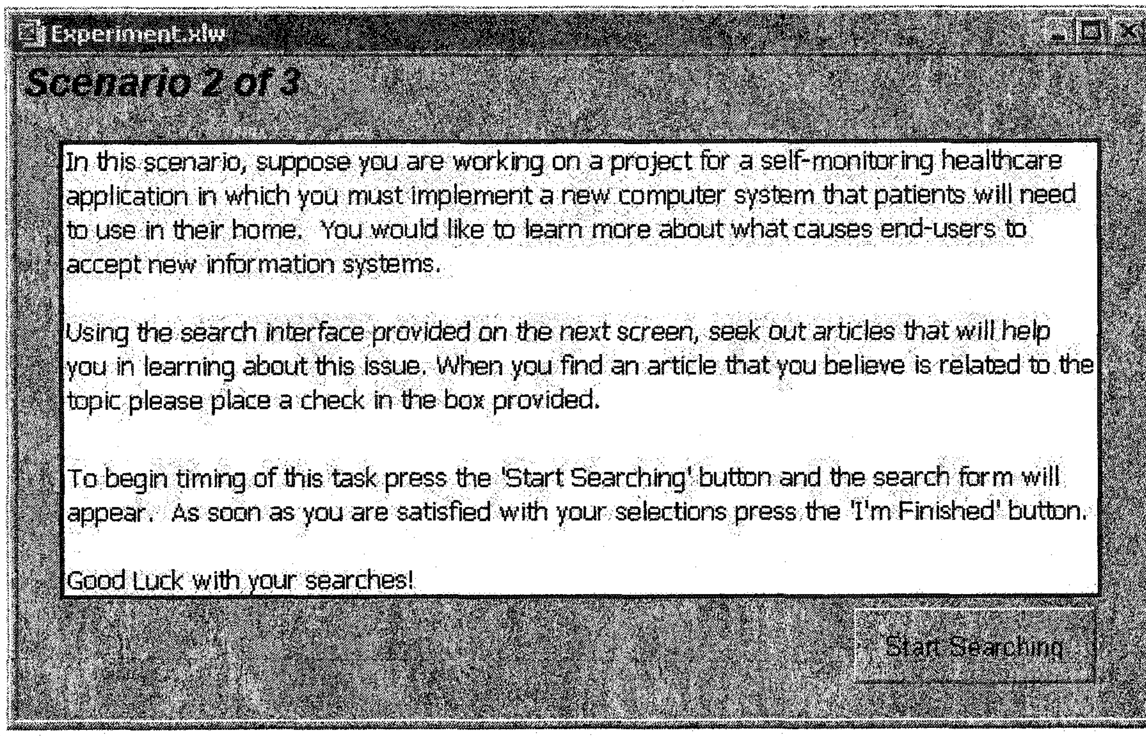

Experiment.xlw
Demographic Data
Please enter the year you were born.
Please enter your gender (Female/Male).
Please select the highest level of education you have attained.
I am considered an IS Professional in my place of work.
I hold an IS (or related) degree.
For the following statements please rate them on a scale from 1 to 7
I perform searches on large computer information systems (ERP, CRM, Data warehouses, etc.).
I perform searches on personal or workgroup size database applications (i.e. Microsoft Access applications, etc.)
I perform searches on the Internet (via Google, Yahoo!, MSN Search, etc.).
<< Back to Introduction
Begin Experiment

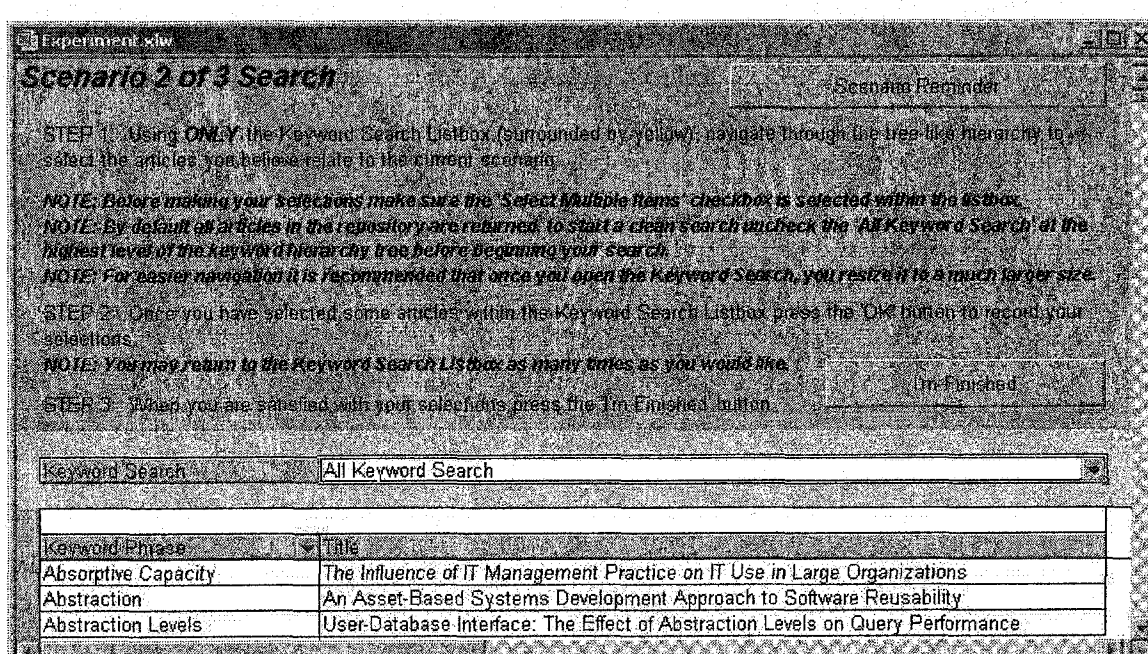

Experiment.xlw
Scenario 1 of 3
In this scenario, suppose you are a manager in a sizable information technology department within a large corporation. Assume your department is responsible for a large number of internal IT application development projects. Many of your projects tend to have problems that you would like to see alleviated. You believe that many of these problems could have been avoided by better design during the information systems development process. Before moving forward with any new projects you would like to learn about ways to more effectively design information systems.
Using the search interface provided on the next screen, seek out articles that will help you in learning about this issue. When you find an article that you believe is related to the topic please place a check in the box provided.
To begin timing of this task press the 'Start Searching' button and the search form will appear. As soon as you are satisfied with your selections press the 'I'm Finished' button.
Good Luck with your searches!
Start Searching

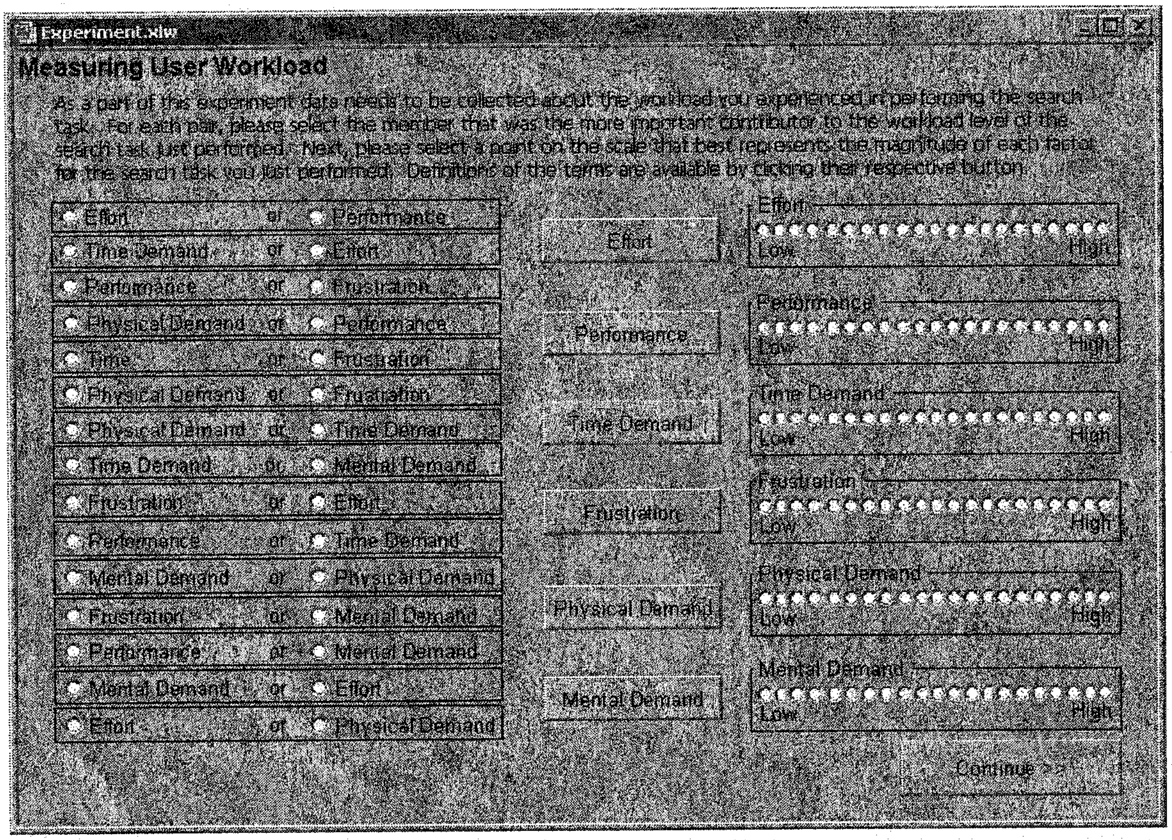

Experiment.xlw
Scenario 1 of 3 Search
Scenario Reminder
STEP 1: Using ONLY the Keyword Search Listbox (surrounded by yellow) navigate through the tree-like hierarchy to select the articles you believe relate to the current scenario.
NOTE: Before making your selections make sure the 'Select Multiple Items' checkbox is selected within the listbox.
NOTE: By default all articles in the repository are returned, to start a clean search uncheck the 'All Keyword Search' at the highest level of the keyword hierarchy tree before beginning your search.
NOTE: For easier navigation it is recommended that once you open the Keyword Search, you resize it to a much larger size.
STEP 2: Once you have selected some articles within the Keyword Search Listbox press the 'OK' button to record your selections.
NOTE: You may return to the Keyword Search Listbox as many times as you would like.
I'm Finished
STEP 3: When you are satisfied with your selections press the 'I'm Finished' button.
Keyword Search
All Keyword Search
Keyword Phrase
Title
Absorptive Capacity
The Influence of IT Management Practice on IT Use in Large Organizations
Abstraction
An Asset-Based Systems Development Approach to Software Reusability
Abstraction Levels
User-Database Interface: The Effect of Abstraction Levels on Query Performance

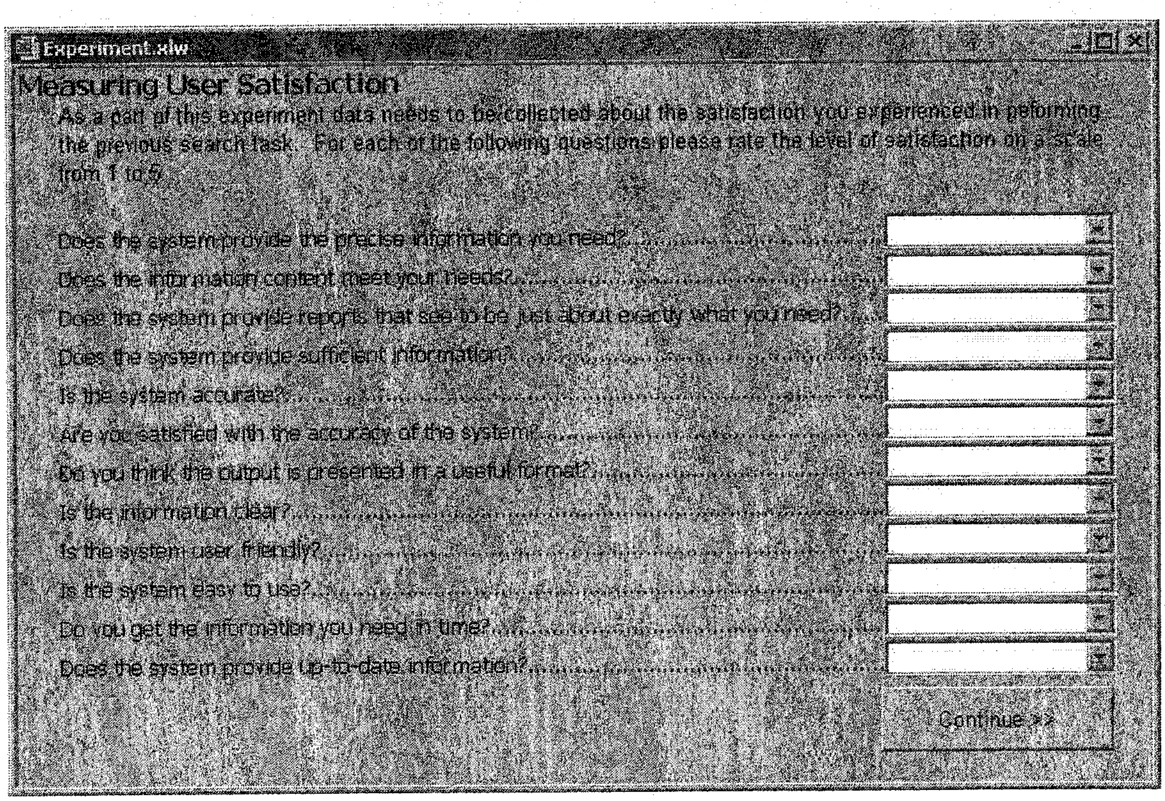

Experiment.xlw
Measuring User Workload
As a part of this experiment data needs to be collected about the workload you experienced in performing the search task. For each pair, please select the member that was the more important contributor to the workload level of the search task just performed. Next, please select a point on the scale that best represents the magnitude of each factor for the search task you just performed. Definitions of the terms are available by clicking their respective button.
Effort or Performance
Time Demand or Effort
Performance or Frustration
Physical Demand or Performance
Time or Frustration
Physical Demand or Frustration
Physical Demand or Time Demand
Time Demand or Mental Demand
Frustration or Effort
Performance or Time Demand
Mental Demand or Physical Demand
Frustration or Mental Demand
Performance or Mental Demand
Mental Demand or Effort
Effort or Physical Demand
Effort
Performance
Time Demand
Frustration
Physical Demand
Mental Demand
Effort Low High
Performance Low High
Time Demand Low High
Frustration Low High
Physical Demand Low High
Mental Demand Low High
Continue >>

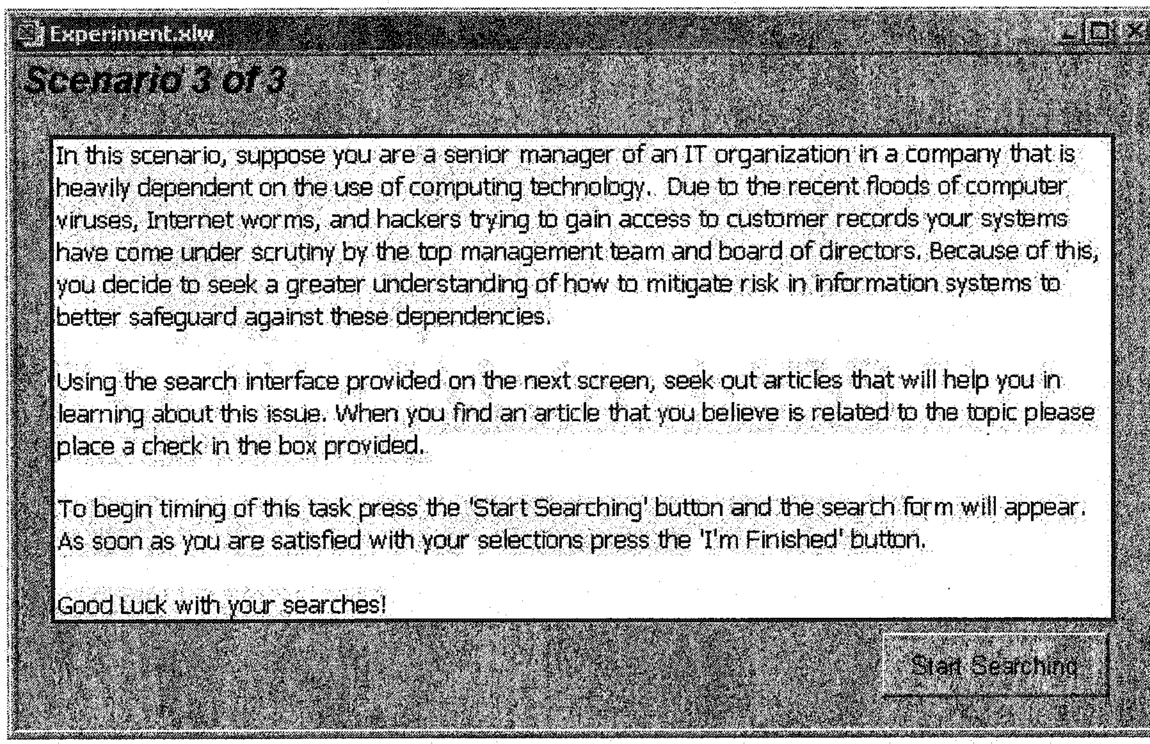

Experiment.xlw
Measuring User Satisfaction
As a part of this experiment data needs to be collected about the satisfaction you experienced in peforming the previous search task. For each of the following questions please rate the level of satisfaction on a scale from 1 to 5.
Does the system provide the precise information you need?
Does the information content meet your needs?
Does the system provide reports that see to be just about exactly what you need?
Does the system provide sufficient information?
Is the system accurate?
Are you satisfied with the accuracy of the system?
Do you think the output is presented in a useful format?
Is the information clear?
Is the system user friendly?
Is the system easy to use?
Do you get the information you need in time?
Does the system provide up-to-date information?
Continue >>

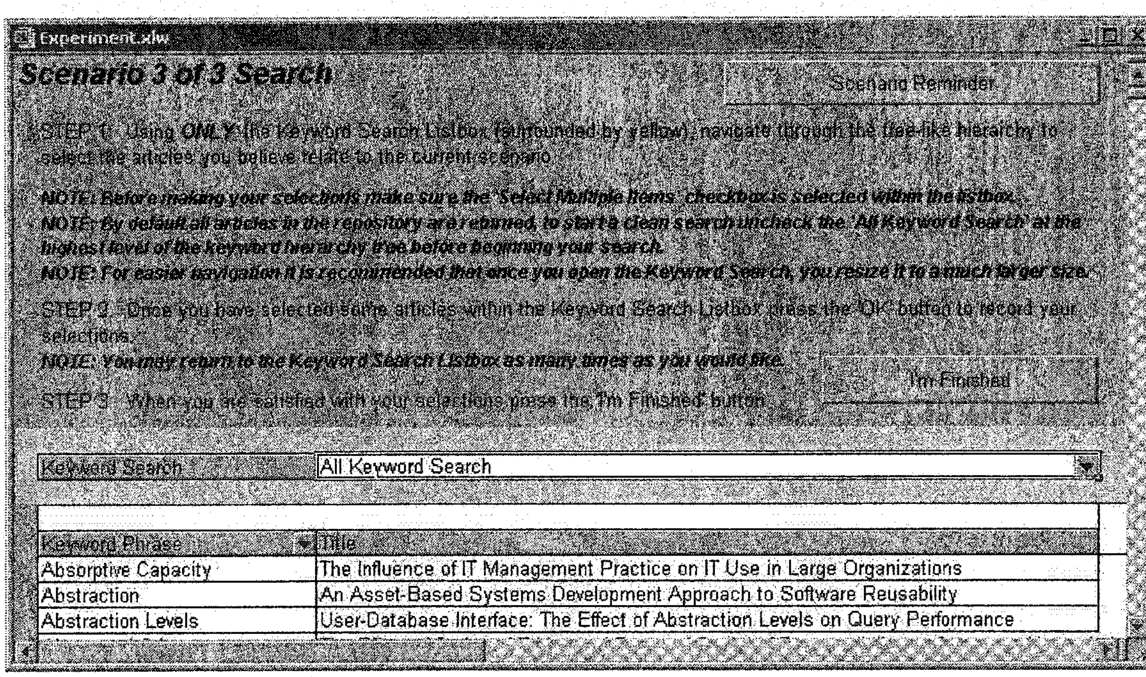

Experiment.xlw
Scenario 2 of 3
In this scenario, suppose you are working on a project for a self-monitoring healthcare application in which you must implement a new computer system that patients will need to use in their home. You would like to learn more about what causes end-users to accept new information systems.
Using the search interface provided on the next screen, seek out articles that will help you in learning about this issue. When you find an article that you believe is related to the topic please place a check in the box provided.
To begin timing of this task press the 'Start Searching' button and the search form will appear. As soon as you are satisfied with your selections press the 'I'm Finished' button.
Good Luck with your searches!
Start Searching

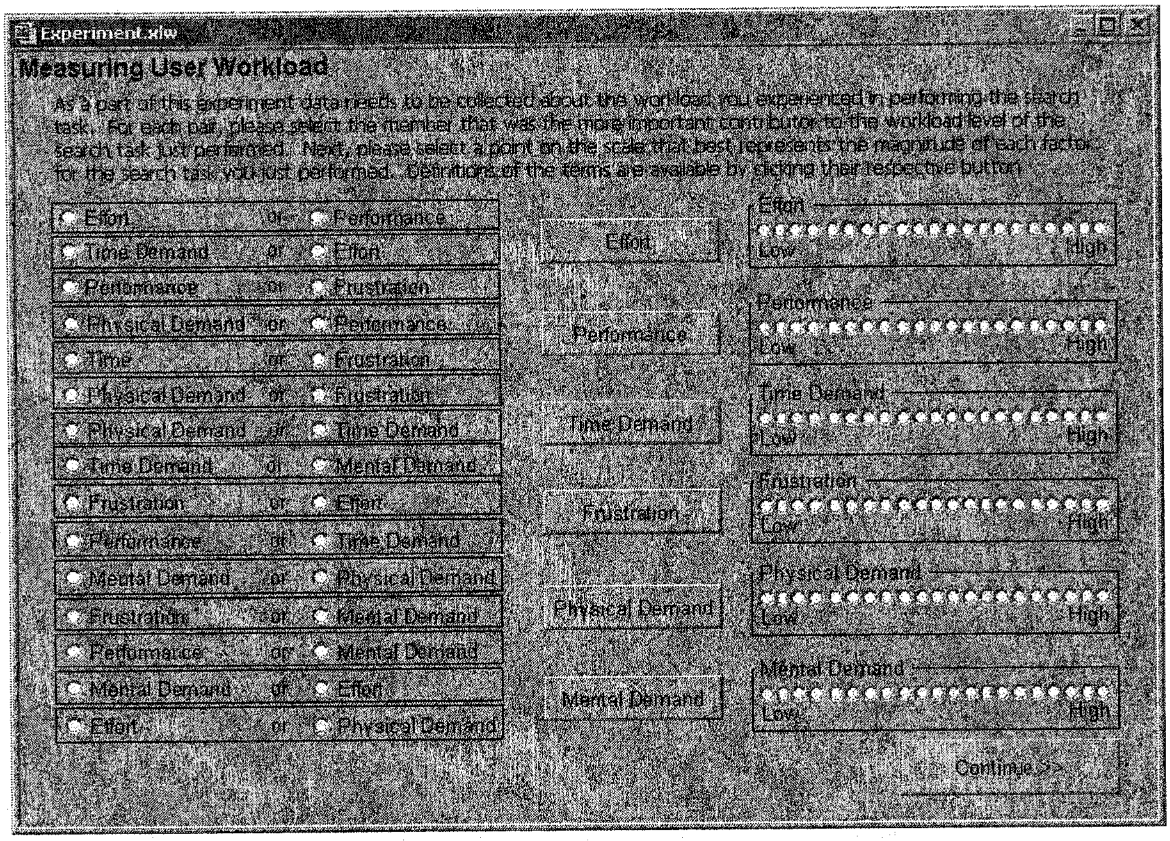

Experiment.xlw
Scenario 2 of 3 Search
Scenario Reminder
STEP 1: Using ONLY the Keyword Search Listbox (surrounded by yellow), navigate through the tree-like hierarchy to select the articles you believe relate to the current scenario.
NOTE: Before making your selections make sure the 'Select Multiple Items' checkbox is selected within the listbox.
NOTE: By default all articles in the repository are returned, to start a clean search uncheck the 'All Keyword Search' at the highest level of the keyword hierarchy tree before beginning your search.
NOTE: For easier navigation it is recommended that once you open the Keyword Search, you resize it to a much larger size.
STEP 2: Once you have selected some articles within the Keyword Search Listbox press the 'OK' button to record your selections.
NOTE: You may return to the Keyword Search Listbox as many times as you would like.
I'm Finished
STEP 3: When you are satisfied with your selections press the 'I'm Finished' button.
Keyword Search
All Keyword Search
Keyword Phrase
Title
Absorptive Capacity
The Influence of IT Management Practice on IT Use in Large Organizations
Abstraction
An Asset-Based Systems Development Approach to Software Reusability
Abstraction Levels
User-Database Interface: The Effect of Abstraction Levels on Query Performance

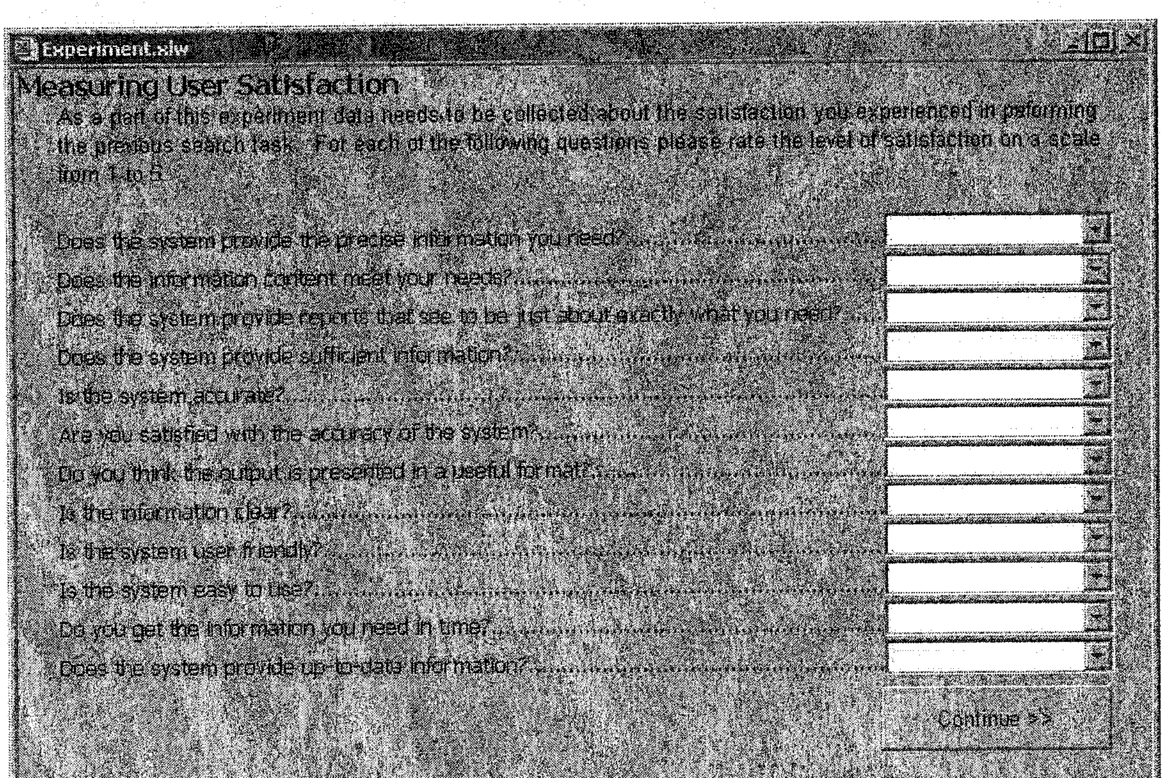

Experiment.xlw
Measuring User Workload
As a part of this experiment data needs to be collected about the workload you experienced in performing the search task. For each pair, please select the member that was the more important contributor to the workload level of the search task just performed. Next, please select a point on the scale that best represents the magnitude of each factor for the search task you just performed. Definitions of the terms are available by clicking their respective button.
Effort or Performance
Time Demand or Effort
Performance or Frustration
Physical Demand or Performance
Time or Frustration
Physical Demand or Frustration
Physical Demand or Time Demand
Time Demand or Mental Demand
Frustration or Effort
Performance or Time Demand
Mental Demand or Physical Demand
Frustration or Mental Demand
Performance or Mental Demand
Mental Demand or Effort
Effort or Physical Demand
Effort
Performance
Time Demand
Frustration
Physical Demand
Mental Demand
Low
High
Continue >>

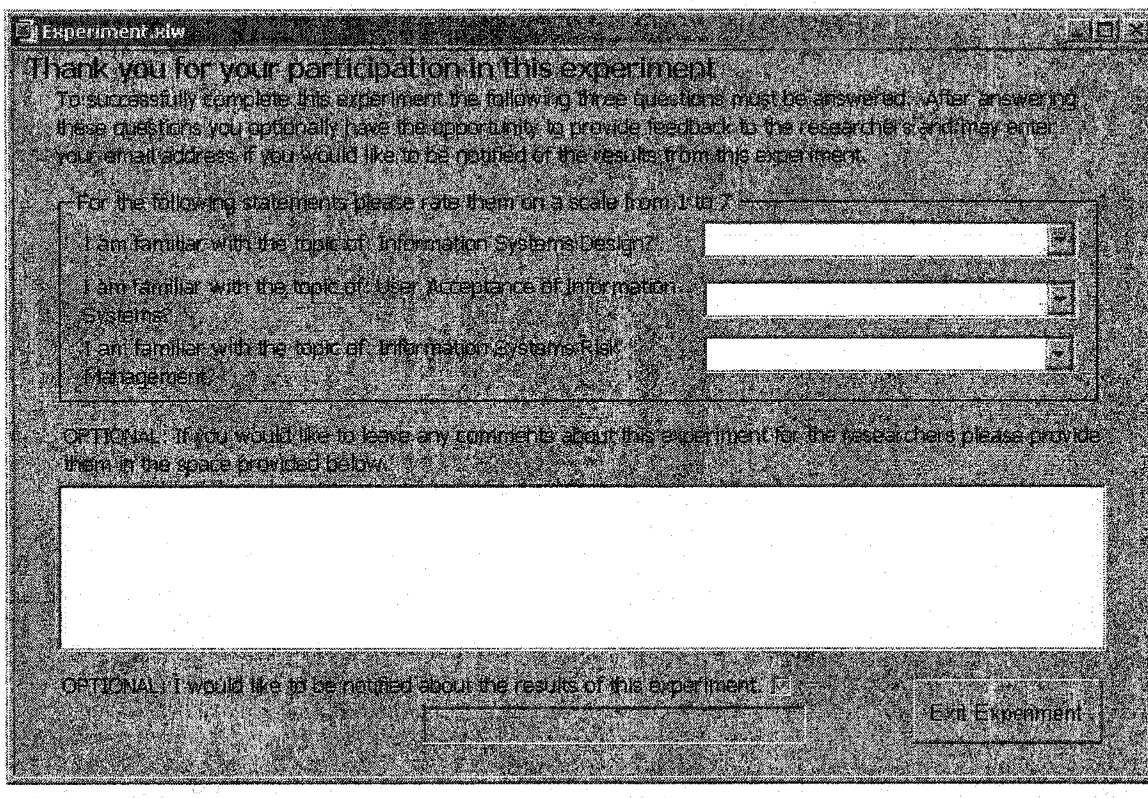

Experiment.xlw
Measuring User Satisfaction
As a part of this experiment data needs to be collected about the satisfaction you experienced in peforming the previous search task. For each of the following questions please rate the level of satisfaction on a scale from 1 to 5.
Does the system provide the precise information you need?
Does the information content meet your needs?
Does the system provide reports that see to be just about exactly what you need?
Does the system provide sufficient information?
Is the system accurate?
Are you satisfied with the accuracy of the system?
Do you think the output is presented in a useful format?
Is the information clear?
Is the system user friendly?
Is the system easy to use?
Do you get the information you need in time?
Does the system provide up-to-date information?
Continue >>


Experiment.xlw
Scenario 3 of 3
In this scenario, suppose you are a senior manager of an IT organization in a company that is heavily dependent on the use of computing technology. Due to the recent floods of computer viruses, Internet worms, and hackers trying to gain access to customer records your systems have come under scrutiny by the top management team and board of directors. Because of this, you decide to seek a greater understanding of how to mitigate risk in information systems to better safeguard against these dependencies.
Using the search interface provided on the next screen, seek out articles that will help you in learning about this issue. When you find an article that you believe is related to the topic please place a check in the box provided.
To begin timing of this task press the 'Start Searching' button and the search form will appear. As soon as you are satisfied with your selections press the 'I'm Finished' button.
Good Luck with your searches!
Start Searching


Experiment.xlw
Scenario 3 of 3 Search
Scenario Reminder
STEP 1: Using ONLY the Keyword Search Listbox (surrounded by yellow), navigate through the tree-like hierarchy to select the articles you believe relate to the current scenario
NOTE: Before making your selections make sure the 'Select Multiple Items' checkbox is selected within the listbox.
NOTE: By default all articles in the repository are returned, to start a clean search uncheck the 'All Keyword Search' at the highest level of the keyword hierarchy tree before beginning your search.
NOTE: For easier navigation it is recommended that once you open the Keyword Search, you resize it to a much larger size.
STEP 2: Once you have selected some articles within the Keyword Search Listbox press the 'OK' button to record your selections.
NOTE: You may return to the Keyword Search Listbox as many times as you would like.
STEP 3: When you are satisfied with your selections press the 'I'm Finished' button.
I'm Finished
Keyword Search
All Keyword Search
Keyword Phrase
Title
Absorptive Capacity
The Influence of IT Management Practice on IT Use in Large Organizations
Abstraction
An Asset-Based Systems Development Approach to Software Reusability
Abstraction Levels
User-Database Interface: The Effect of Abstraction Levels on Query Performance


Experiment.xlw
Measuring User Workload
As a part of this experiment data needs to be collected about the workload you experienced in performing the search task. For each pair, please select the member that was the more important contributor to the workload level of the search task just performed. Next, please select a point on the scale that best represents the magnitude of each factor for the search task you just performed. Definitions of the terms are available by clicking their respective button.
Effort or Performance
Time Demand or Effort
Performance or Frustration
Physical Demand or Performance
Time or Frustration
Physical Demand or Frustration
Physical Demand or Time Demand
Time Demand or Mental Demand
Frustration or Effort
Performance or Time Demand
Mental Demand or Physical Demand
Frustration or Mental Demand
Performance or Mental Demand
Mental Demand or Effort
Effort or Physical Demand
Effort
Performance
Time Demand
Frustration
Physical Demand
Mental Demand
Effort
Low High
Performance
Low High
Time Demand
Low High
Frustration
Low High
Physical Demand
Low High
Mental Demand
Low High
Continue >>


Experiment.xlw
Measuring User Satisfaction
As a part of this experiment data needs to be collected about the satisfaction you experienced in peforming the previous search task. For each of the following questions please rate the level of satisfaction on a scale from 1 to 5.
Does the system provide the precise information you need?
Does the information content meet your needs?
Does the system provide reports that see to be just about exactly what you need?
Does the system provide sufficient information?
Is the system accurate?
Are you satisfied with the accuracy of the system?
Do you think the output is presented in a useful format?
Is the information clear?
Is the system user friendly?
Is the system easy to use?
Do you get the information you need in time?
Does the system provide up-to-date information?
Continue >>


Experiment.xlw

Thank you for your participation in this experiment

To successfully complete this experiment the following three questions must be answered. After answering these questions you optionally have the opportunity to provide feedback to the researchers and may enter your email address if you would like to be notified of the results from this experiment.

For the following statements please rate them on a scale from 1 to 7

I am familiar with the topic of: Information Systems Design?

I am familiar with the topic of: User Acceptance of Information Systems?

I am familiar with the topic of: Information Systems Risk Management?

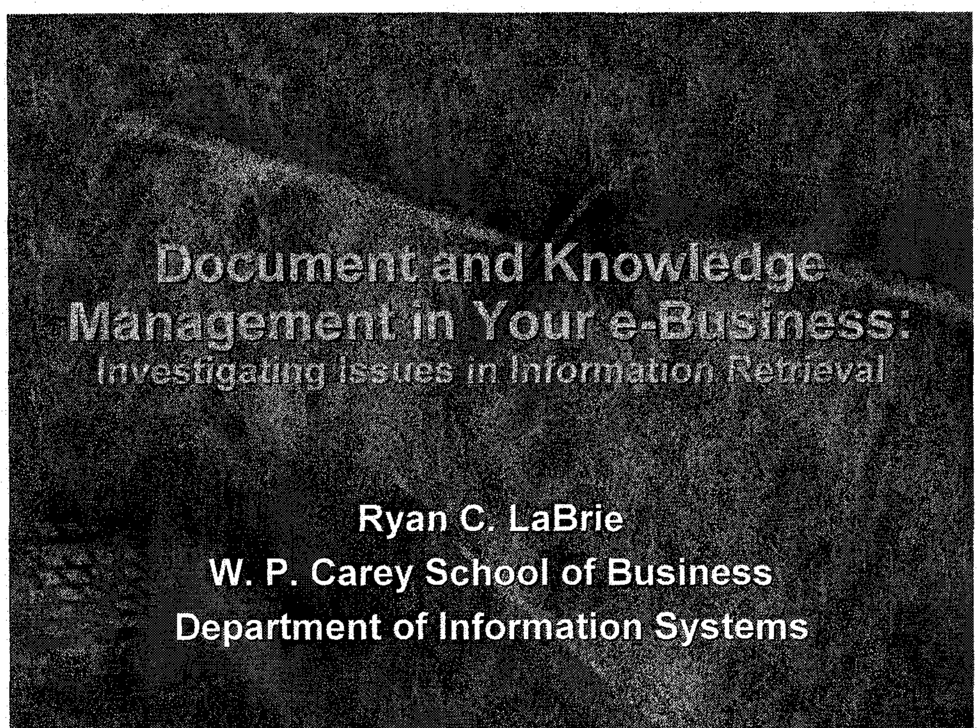
OPTIONAL: If you would like to leave any comments about this experiment for the researchers please provide them in the space provided below.

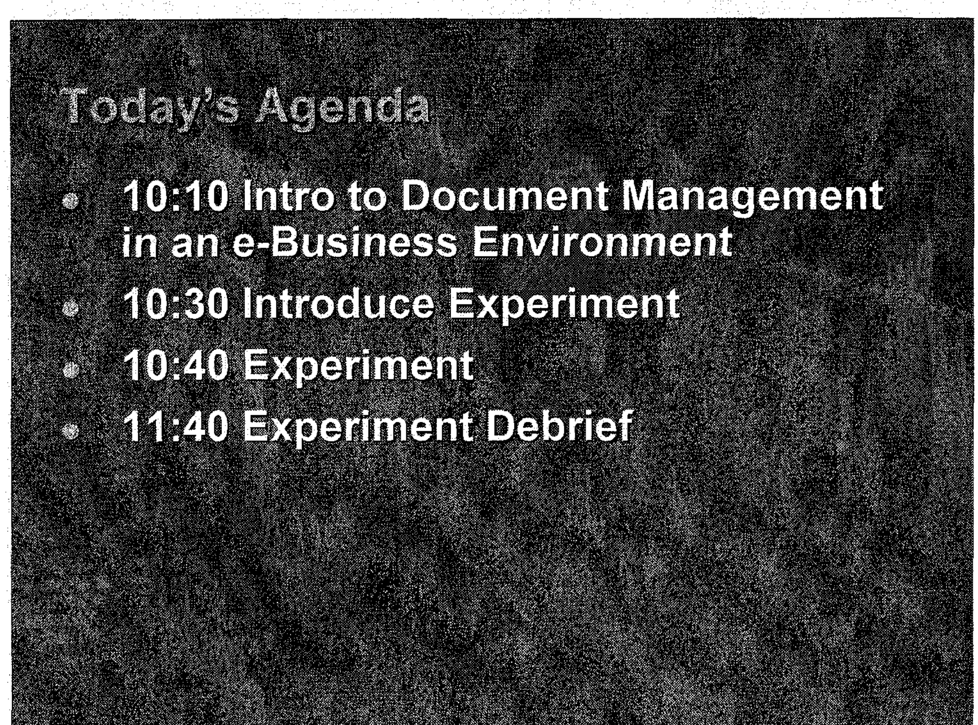
OPTIONAL: I would like to be notified about the results of this experiment.

Exit Experiment

APPENDIX C

EXPERIMENT INTRODUCTION SLIDE DECK


Document and Knowledge
Management in Your e-Business:
Investigating Issues in Information Retrieval
Ryan C. LaBrie
W. P. Carey School of Business
Department of Information Systems


Today's Agenda
10:10 Intro to Document Management in an e-Business Environment
10:30 Introduce Experiment
10:40 Experiment
11:40 Experiment Debrief

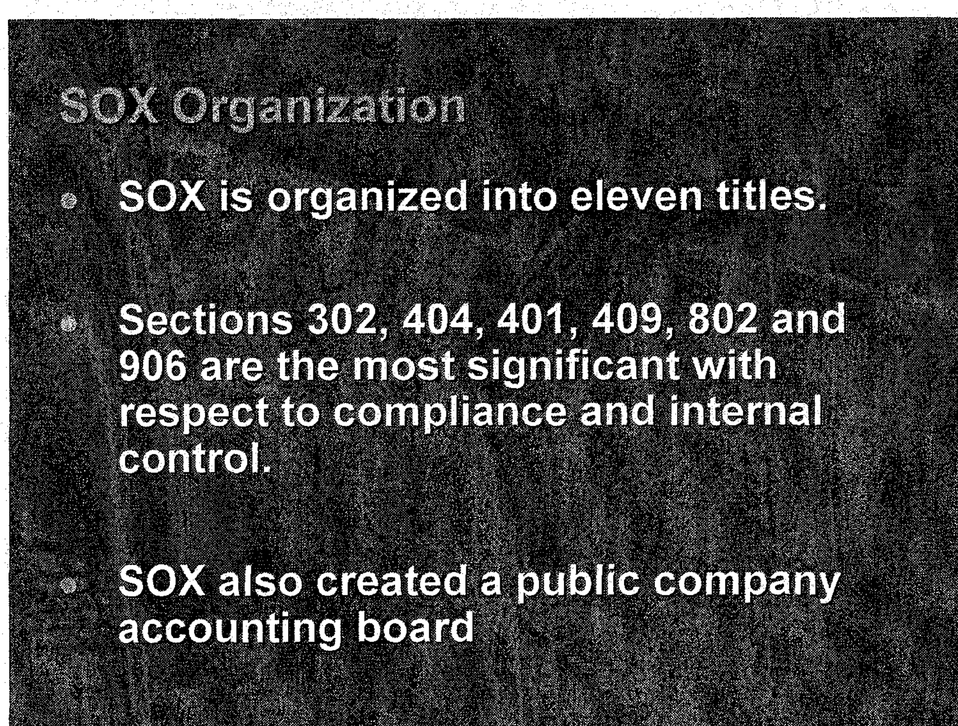
Intangible Assets

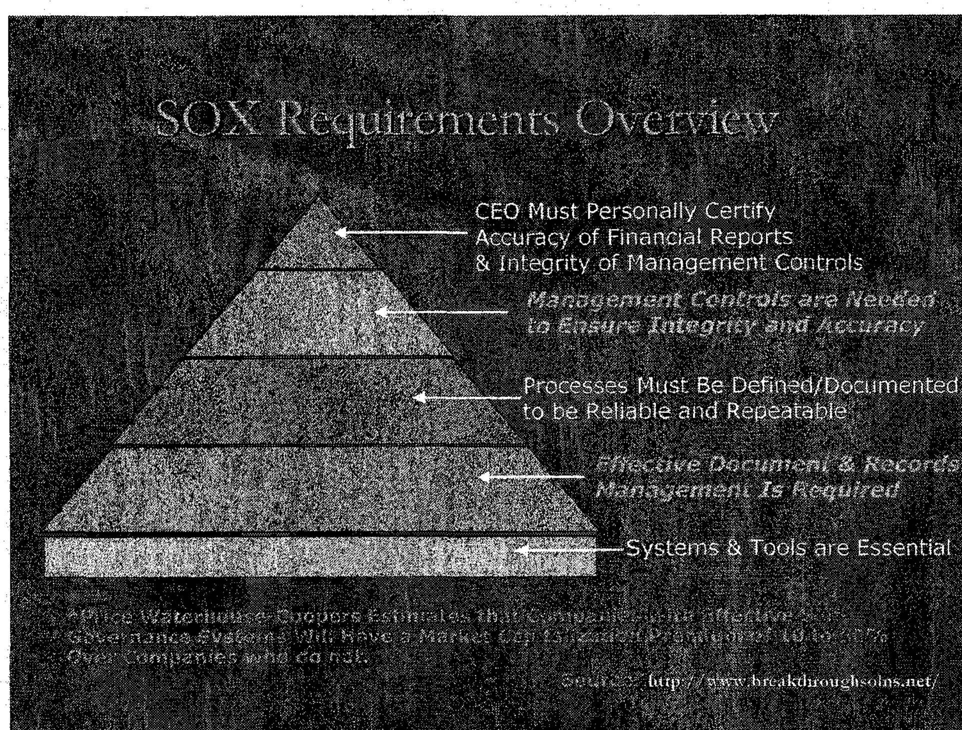
- Alan Greenspan has noted that 75% of organizational assets are now knowledge assets
- The 500 largest firms in the United States had intangible assets valued at $7.3 trillion (70% of their total value)
  - These 500 companies employ more than 21.6 million employees and generate over US$ 6.1 trillion in revenue (Stone, 2002)
- An Arthur Anderson study revealed
  - In 1978 the balance sheet explained 95% of the market value of the firms.
  - In 1998 the balance sheet only explained 28% of the market value of those firms.
- Today, the balance sheet explains less than 15% of the market value of the average firm (Stanfield, 2002)

What is the Sarbanes-Oxley Act (SOX)?

- Named after its main architects, Senator Paul Sarbanes and Representative Michael Oxley
- The Sarbanes-Oxley Act was signed into law on 30 July 2002
- Most public companies must meet the financial reporting and certification mandates for any end of year financial statements filed after 15 June 2004 (everybody by 15 April 2005)
- Introduces highly significant legislative changes to financial practice and corporate governance regulation
  - Aggressive accounting and earnings management
  - Abuses of stock options compensations
  - Significant restatements
  - Erosion of ethics


Why Do We Need SOX?
Spectacular Corporate, Governance, and Audit Failures:
ENRON
WORLDCOM
GLOBAL CROSSING
TYCO
ADELPHIA
...AND MANY OTHERS


SOX Organization
SOX is organized into eleven titles.
Sections 302, 404, 401, 409, 802 and 906 are the most significant with respect to compliance and internal control.
SOX also created a public company accounting board


SOX Requirements Overview
CEO Must Personally Certify Accuracy of Financial Reports & Integrity of Management Controls
Management Controls are Needed to Ensure Integrity and Accuracy
Processes Must Be Defined/Documented to be Reliable and Repeatable
Effective Document & Records Management Is Required
Systems & Tools are Essential
Over Companies who do not.
http://www.breakthroughsolns.net/

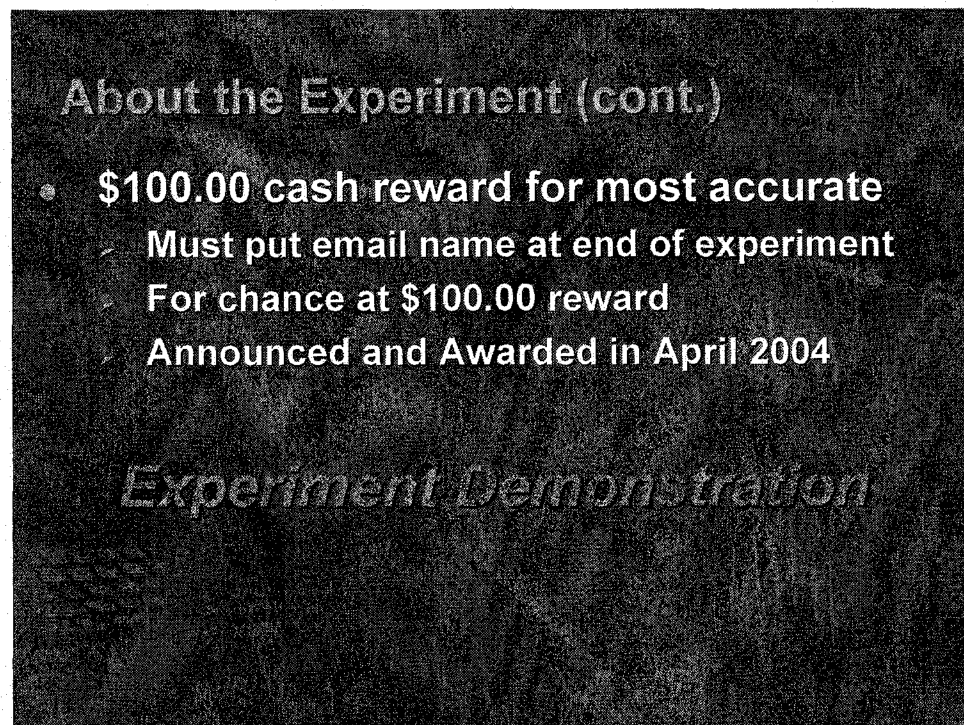

IT role in SOX compliance
Communication and Reporting
Documentation
Internal Control Assessments
Measurements
Monitoring

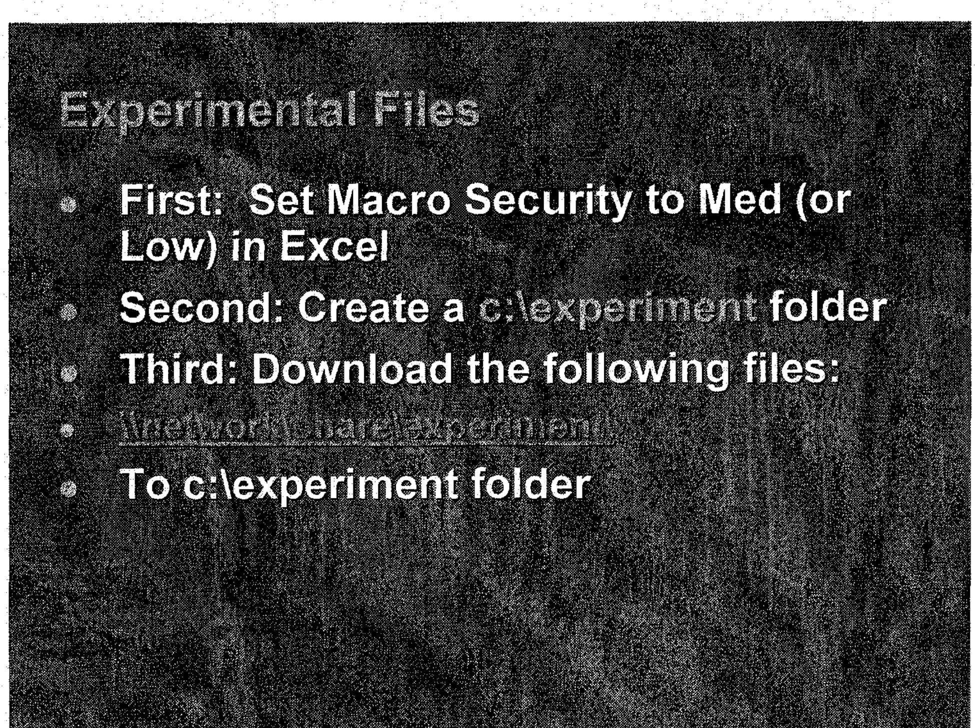
IT Role: Documentation

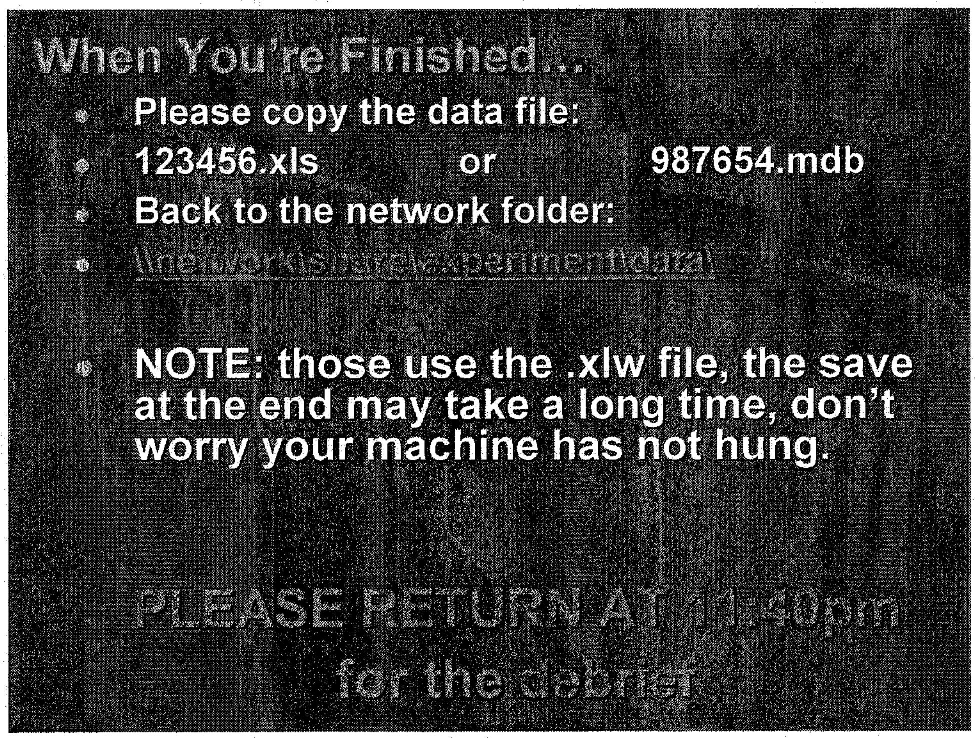
- Need a centralized system to document their internal control environment.
- Management and Process/Control owners across the organization should have *anytime and anywhere access* to documentation elements
- Technology solutions exist to centrally create and *manage digital documents* allowing worldwide access via the corporate intranets with a single authentication and access control security

About the Experiment

- Tests two (2) information retrieval (search) mechanisms
- Document management retrieval context
- Each individual will perform three (3) searches in only one search interface
- Collects data on:
  - Experience
  - Accuracy
  - Timeliness
  - Work Effort
  - Satisfaction


About the Experiment (cont.)
$100.00 cash reward for most accurate
Must put email name at end of experiment
For chance at $100.00 reward
Announced and Awarded in April 2004
Experiment Demonstration


Experimental Files
First: Set Macro Security to Med (or Low) in Excel
Second: Create a c:\experiment folder
Third: Download the following files:
\\network\share\experiment\
To c:\experiment folder


When You're Finished...
Please copy the data file:
123456.xls or 987654.mdb
Back to the network folder:
\\network\share\experimentdata\
NOTE: those use the .xlw file, the save at the end may take a long time, don't worry your machine has not hung.
PLEASE RETURN AT 11:40pm
for the debrief

APPENDIX D

EXPERIMENT DEBRIEF SLIDE DECK

Experiment Debrief
What is Knowledge Management Research?

Ryan C. LaBrie
W. P. Carey School of Business
Department of Information Systems

About My Dissertation

"The Impact of Alternative Search Mechanisms on the Effectiveness of Knowledge Retrieval"

- Theoretical Basis
- Research Model
- Research Method
  - Statistical Analysis
  - How the experiment was developed
- Implications

Theory

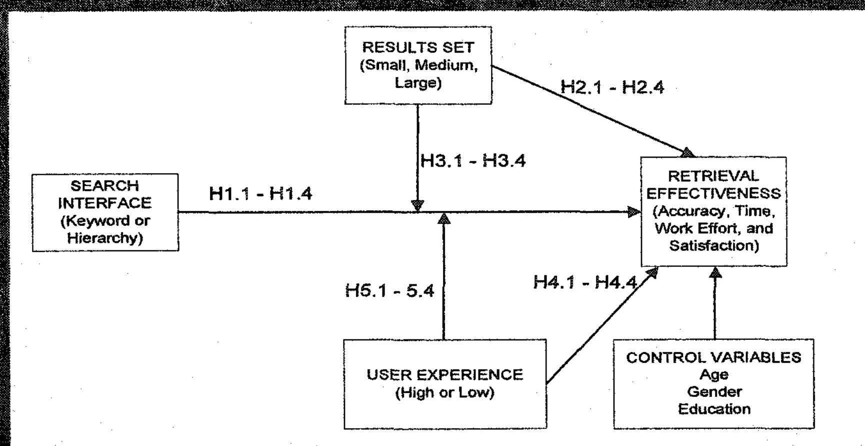
- Domain: Information System (Knowledge Management)
- Reference Discipline
  - Cognitive Psychology
- RECALL vs. RECOGNITION
- Limitation of Keyword Searches
  - Information Overload
  - Ambiguity

Keyword Ambiguity Issue

Keyword Phrase	# of Articles
System Design (original query)	3
Systems Design	10
System Designs	0
Systems Designs	1
System Analysis and Design	0
Systems Analysis and Design	4
Information System Design	2
Information Systems Design	8
MIS Systems Design	1
Participative System Design	1
System Design Methods	1
Expert Systems Design	1
Impact and Socio-Technical Systems Design	1
TOTAL	34

- Gorla & Walker (1998)
  - Ambiguity biggest problem
  - Call for a controlled vocabulary
- LaBrie & St. Louis (2003)
  - Keyword search limitations
  - Mirrored results in classification scheme usage

Research Model

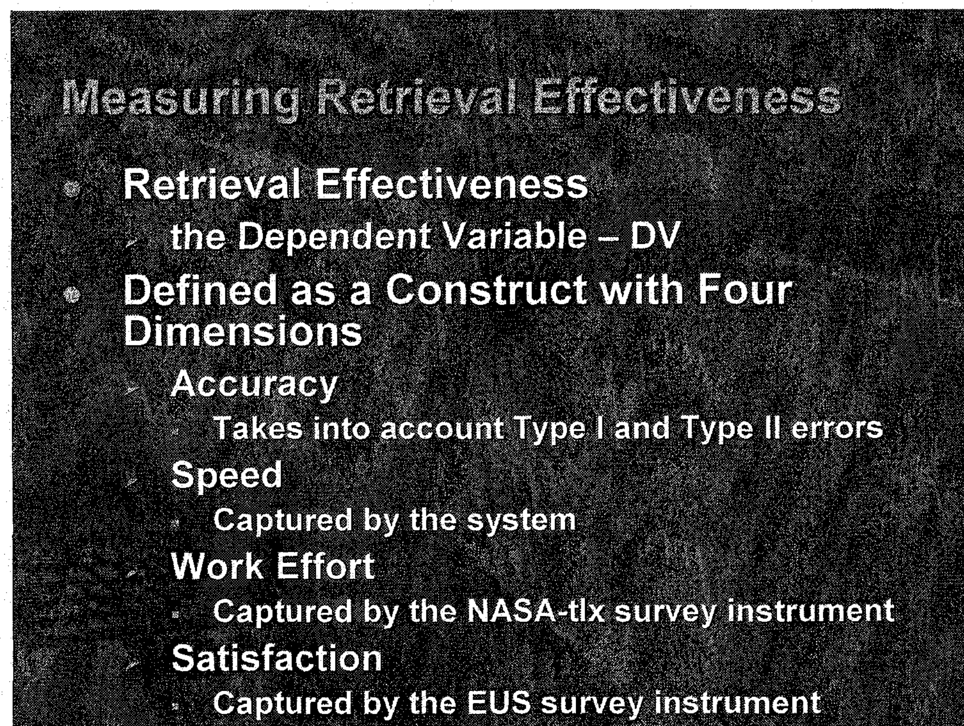


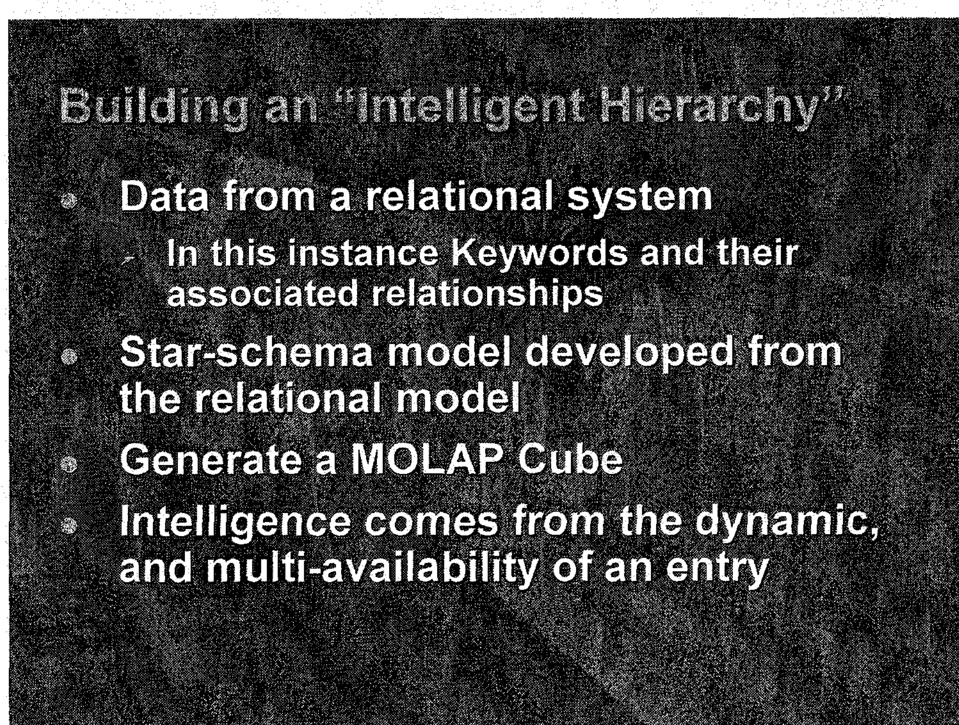
$$RE = \beta 0 + \beta 1*SI + \beta 2*RS + \beta 3*UE + \beta 4*(SI*RS) + \beta 5*(SI*UE) + \epsilon$$

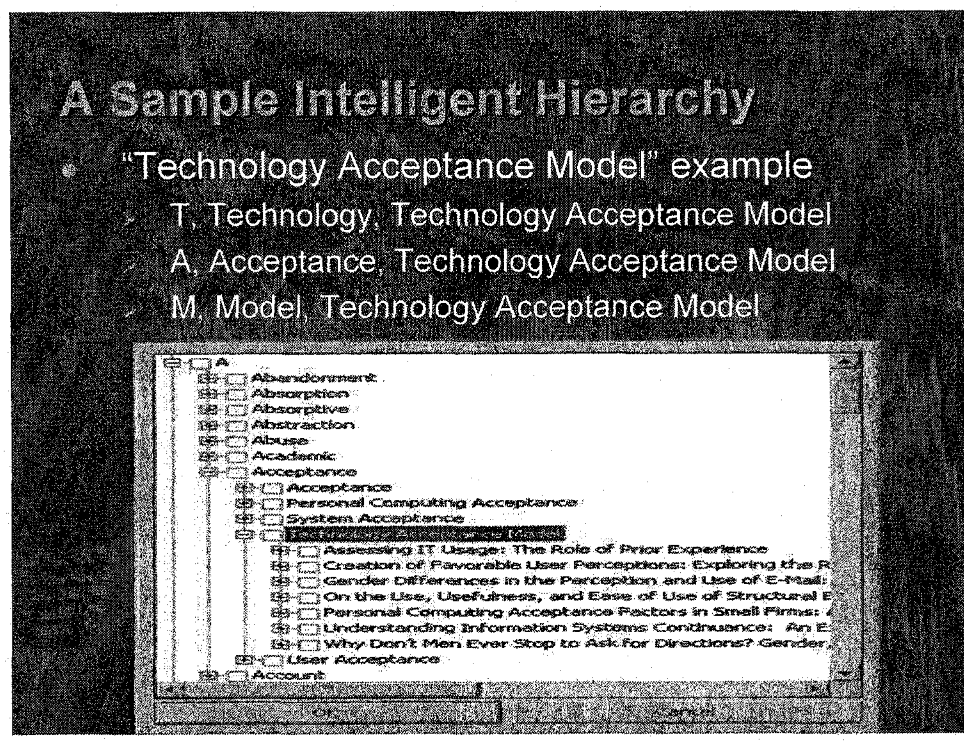
RE = Retrieval Effectiveness   SI = Search Interface
RS = Results Set   UE = User Experience

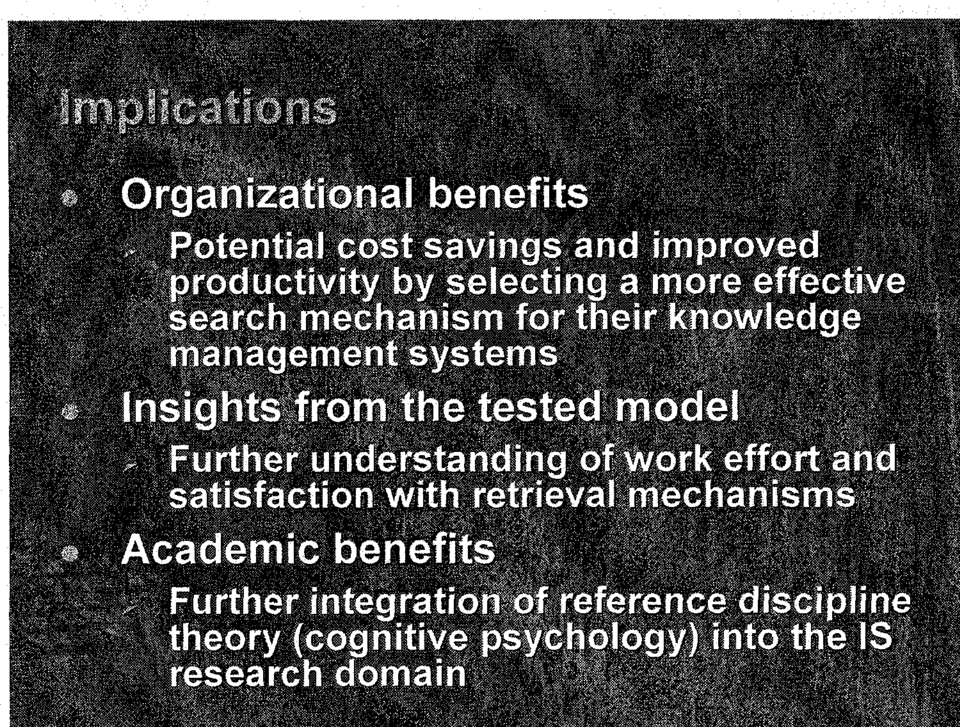
Research Method

- A laboratory experiment
  - vs. Case Study, Survey, Modeling & Simulation, Secondary Data Analysis
- Specifically a: 2 x 3 x 2 mixed factor, analysis of covariance (ANCOVA), non-repeated measures, experimental design
- Factors (Experimental, Independent Variables)
  - *Search Type (keyword or hierarchy – between subjects)*
  - Results Set (small, medium, and large – within subjects)
  - User Experience (low or high – between subjects)


Measuring Retrieval Effectiveness
Retrieval Effectiveness
the Dependent Variable – DV
Defined as a Construct with Four Dimensions
Accuracy
Takes into account Type I and Type II errors
Speed
Captured by the system
Work Effort
Captured by the NASA-tlx survey instrument
Satisfaction
Captured by the EUS survey instrument


Building an "Intelligent Hierarchy"
Data from a relational system
In this instance Keywords and their associated relationships
Star-schema model developed from the relational model
Generate a MOLAP Cube
Intelligence comes from the dynamic, and multi-availability of an entry


A Sample Intelligent Hierarchy
"Technology Acceptance Model" example
T, Technology, Technology Acceptance Model
A, Acceptance, Technology Acceptance Model
M, Model, Technology Acceptance Model
A
Abandonment
Absorption
Absorptive
Abstraction
Abuse
Academic
Acceptance
Acceptance
Personal Computing Acceptance
System Acceptance
Technology Acceptance Model
Assessing IT Usage: The Role of Prior Experience
Creation of Favorable User Perceptions: Exploring the
Gender Differences in the Perception and Use of E-Mail:
On the Use, Usefulness, and Ease of Use of Structural
Personal Computing Acceptance Factors in Small Firms:
Understanding Information Systems Continuance: An
Why Don't Men Ever Stop to Ask for Directions? Gender,
User Acceptance
Account


Implications
Organizational benefits
Potential cost savings and improved productivity by selecting a more effective search mechanism for their knowledge management systems
Insights from the tested model
Further understanding of work effort and satisfaction with retrieval mechanisms
Academic benefits
Further integration of reference discipline theory (cognitive psychology) into the IS research domain

BIOGRAPHICAL SKETCH

Ryan Craig LaBrie was born in Seattle, Washington, United States of America, in 1968. He graduated from Sumner High School in Sumner, Washington in 1986. Ryan graduated with a Bachelors of Science degree in Computer Science (Systems Emphasis) in 1993 from Seattle Pacific University in Seattle, Washington. In 1997, Ryan graduated with a Masters of Science degree in Information Systems Management from Seattle Pacific University. A large portion of Ryan's Masters Thesis project was published as chapters in the books entitled, *Access 95 Power Programming* and *Access 97 Power Programming*. Following a ten year (1990-2000) career with the Microsoft Corporation, where he held the following positions: Systems Support Engineer, Training Specialist, Technical Team Manager, and Program Manager, Ryan began his doctoral studies at Arizona State University. In August 2004 Ryan earned his Doctor of Philosophy in Business Administration (Information Systems) from the W. P. Carey School of Business. Ryan's research interests include: knowledge management, information ethics, and international IT issues. Ryan has taught at the university level (both undergraduate and graduate courses) since 1997. His teaching interests include: database, distributed computing, electronic commerce, information ethics, and networking. To date, Ryan has published articles in the *Annals of Cases on Information Technology*, the *Journal of Electronic Commerce Research,* a number of national and international conferences, and doctoral consortiums. In 2004, upon earning his doctoral degree, Ryan accepted an Associate Professor position with his alma mater, Seattle Pacific University. Ryan resides with his wife, Jennifer and their two children, Rachel and Alexander, in Mercer Island, Washington.